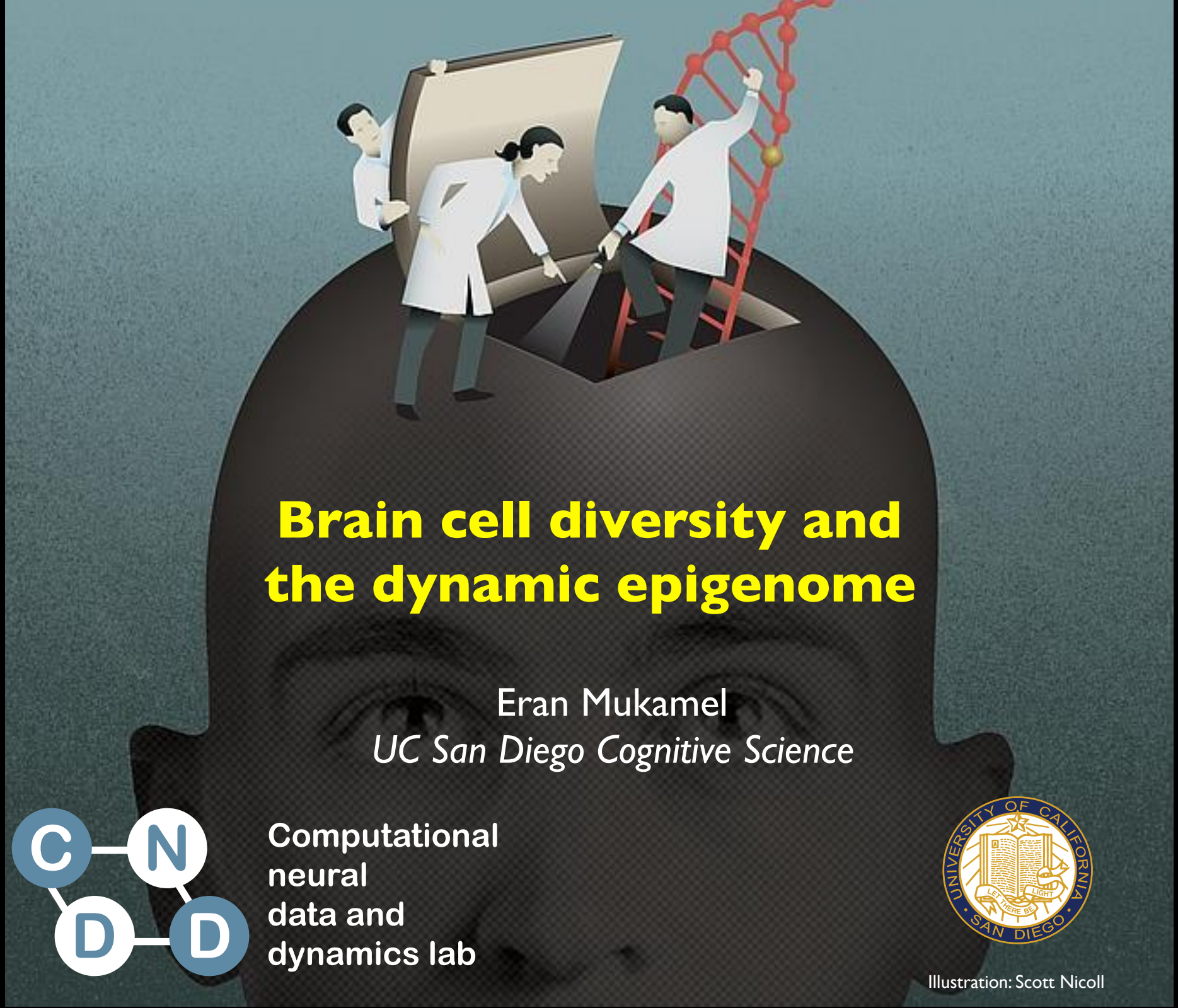
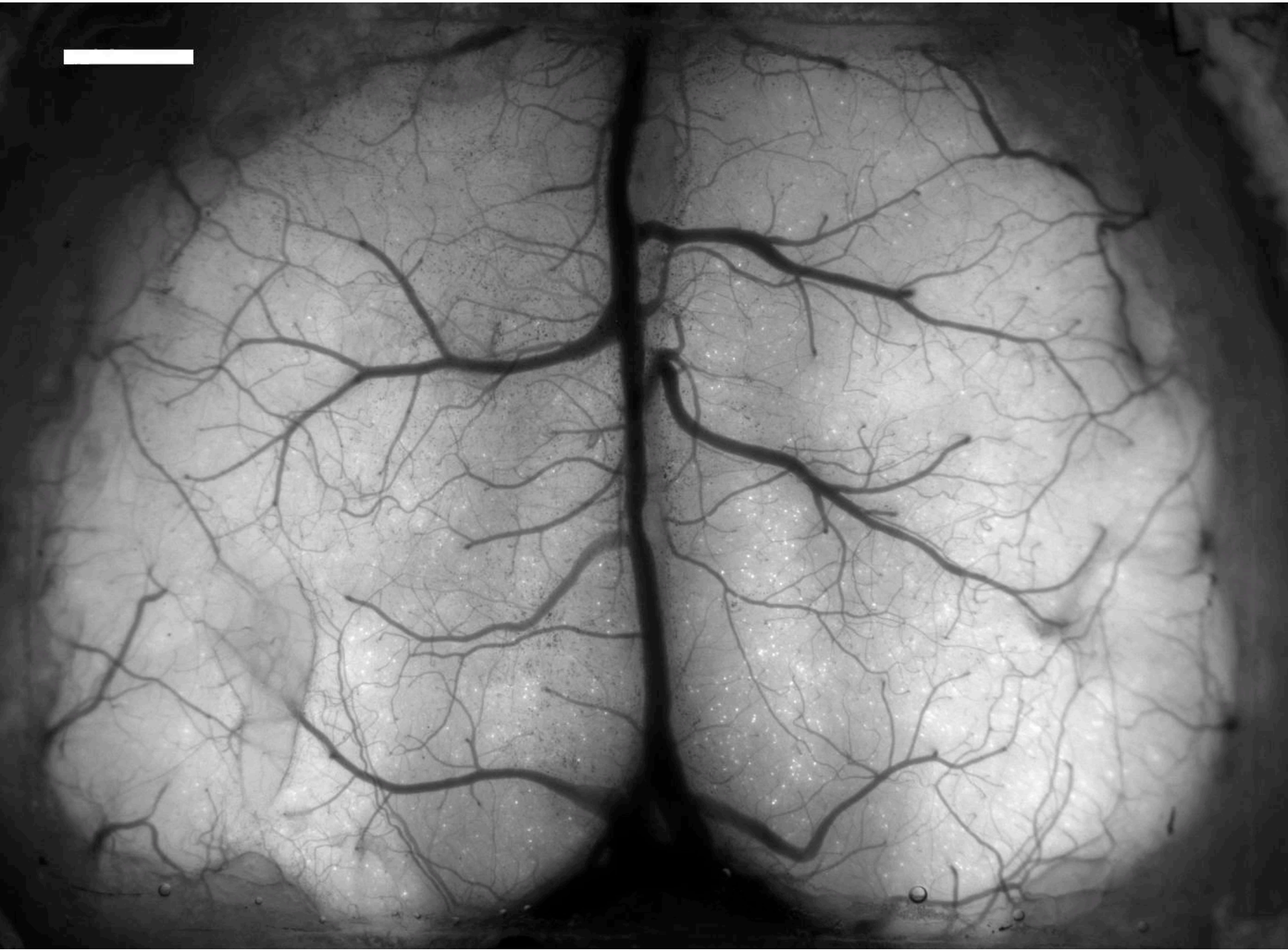


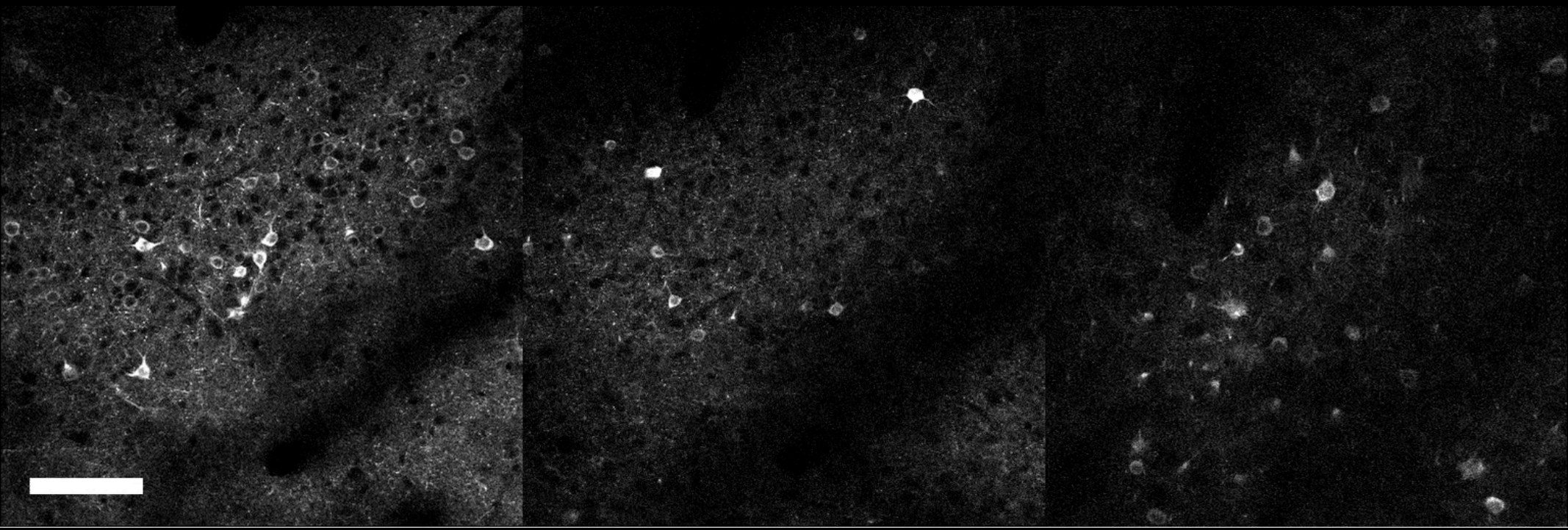
Brain cell diversity and the dynamic epigenome

Eran Mukamel
UC San Diego Cognitive Science

Computational
neural
data and
dynamics lab

Illustration: Scott Nicoll





THE NEW YORKER

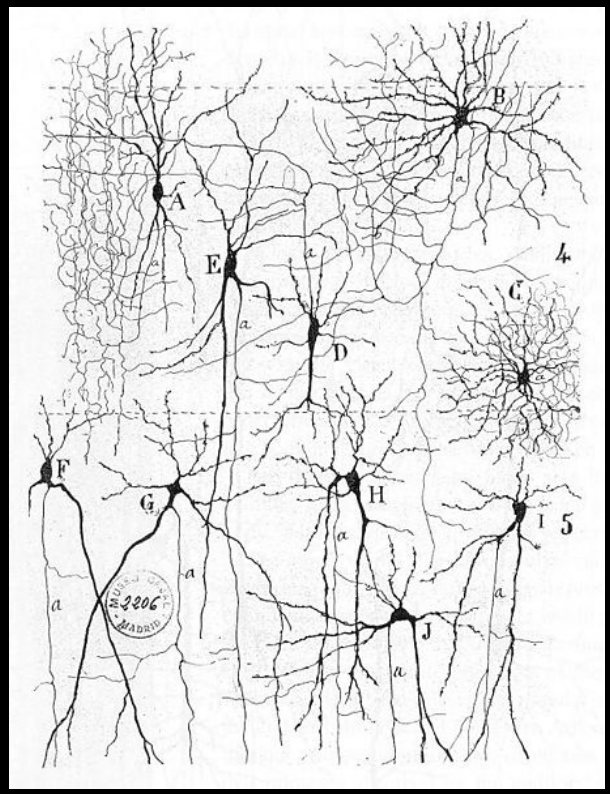
THE MIND-EXPANDING IDEAS OF ANDY CLARK

The tools we use to help us think—from language to smartphones—may be part of thought itself.

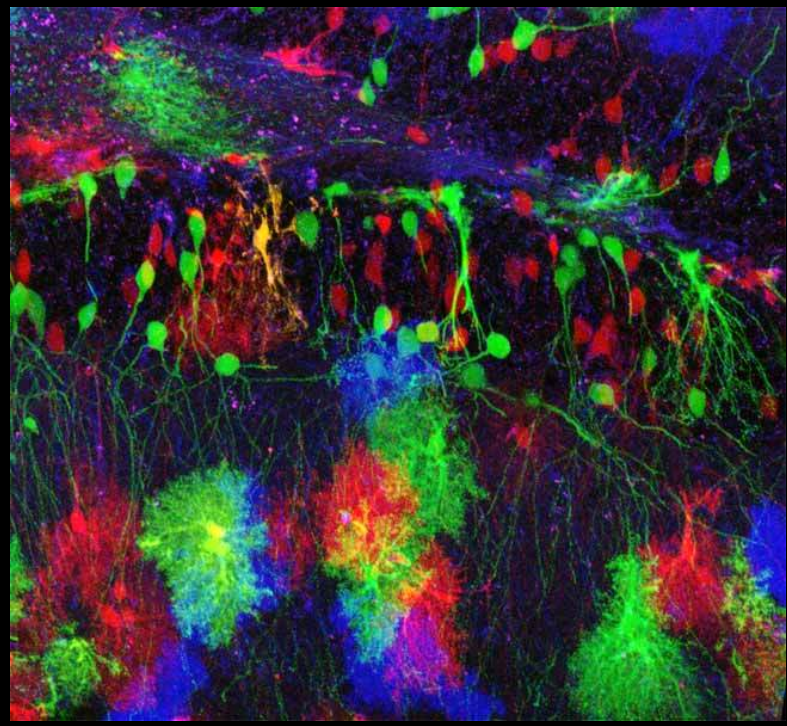
By Larissa MacFarquhar

"There's something very interesting about life," Clark says, "which is that we do seem to be built of system upon system upon system. **The smallest systems are the individual cells, which have an awful lot of their own little intelligence, if you like—they take care of themselves, they have their own things to do.** Maybe there's a great flexibility in being built out of all these little bits of stuff that have their own capacities to protect and organize themselves. I've become more and more open to the idea that some of the fundamental features of life really *are* important to understanding how our mind is possible. I didn't use to think that. I used to think that you could start about halfway up and get everything you needed."

Not all neurons are created equal:



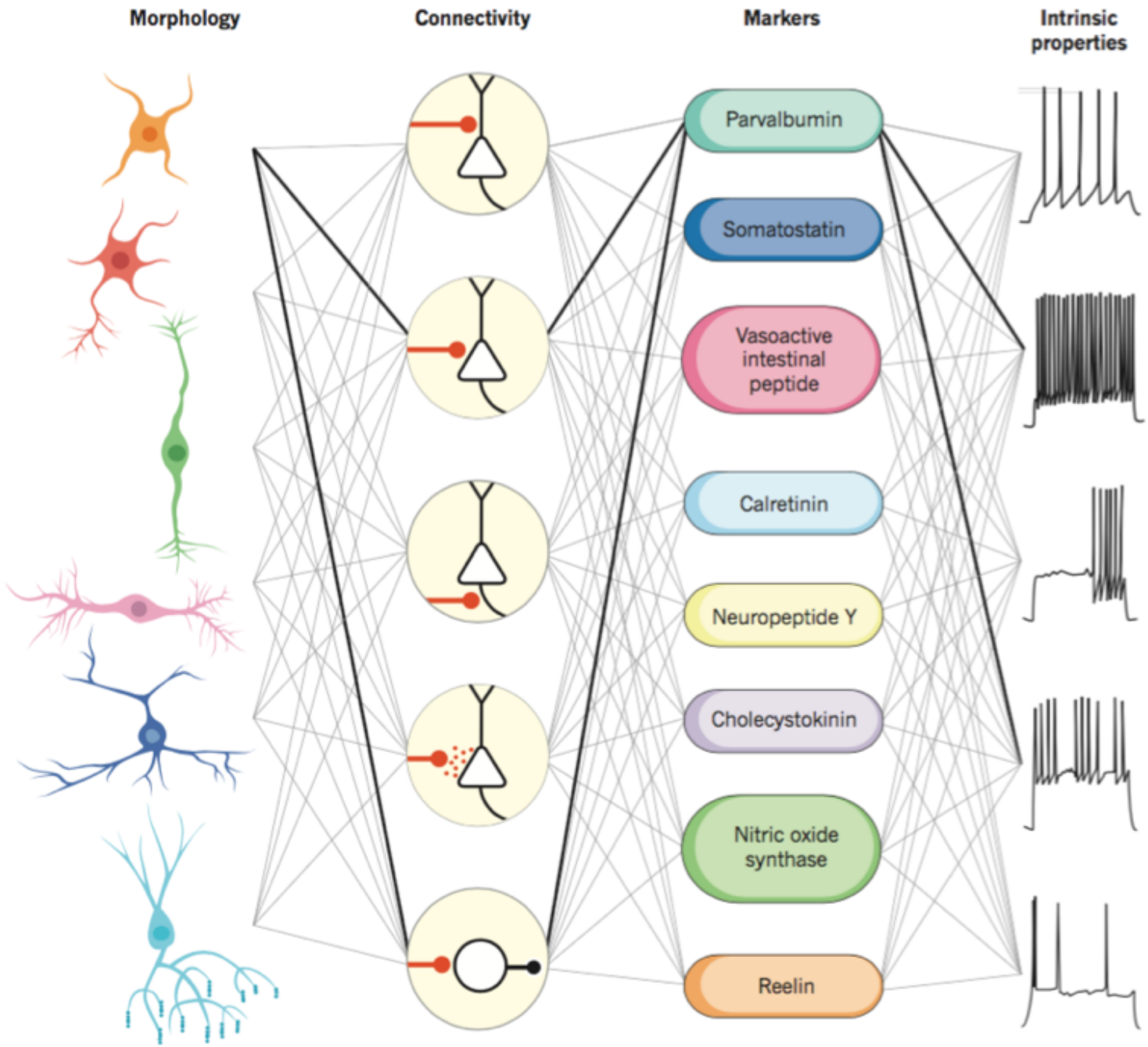
Drawing of auditory cortical neurons, Ramón y Cajal (1899)



Fluorescently labeled neurons and glia,
Livet, Sanes, and Lichtman (2007)

Each neuron has a unique:

1. Location (area, layer)
2. Connections (inputs, outputs)
3. Electrical and chemical responses



Kepecs, A., & Fishell, G. (2014)

SAME BUT DIFFERENT

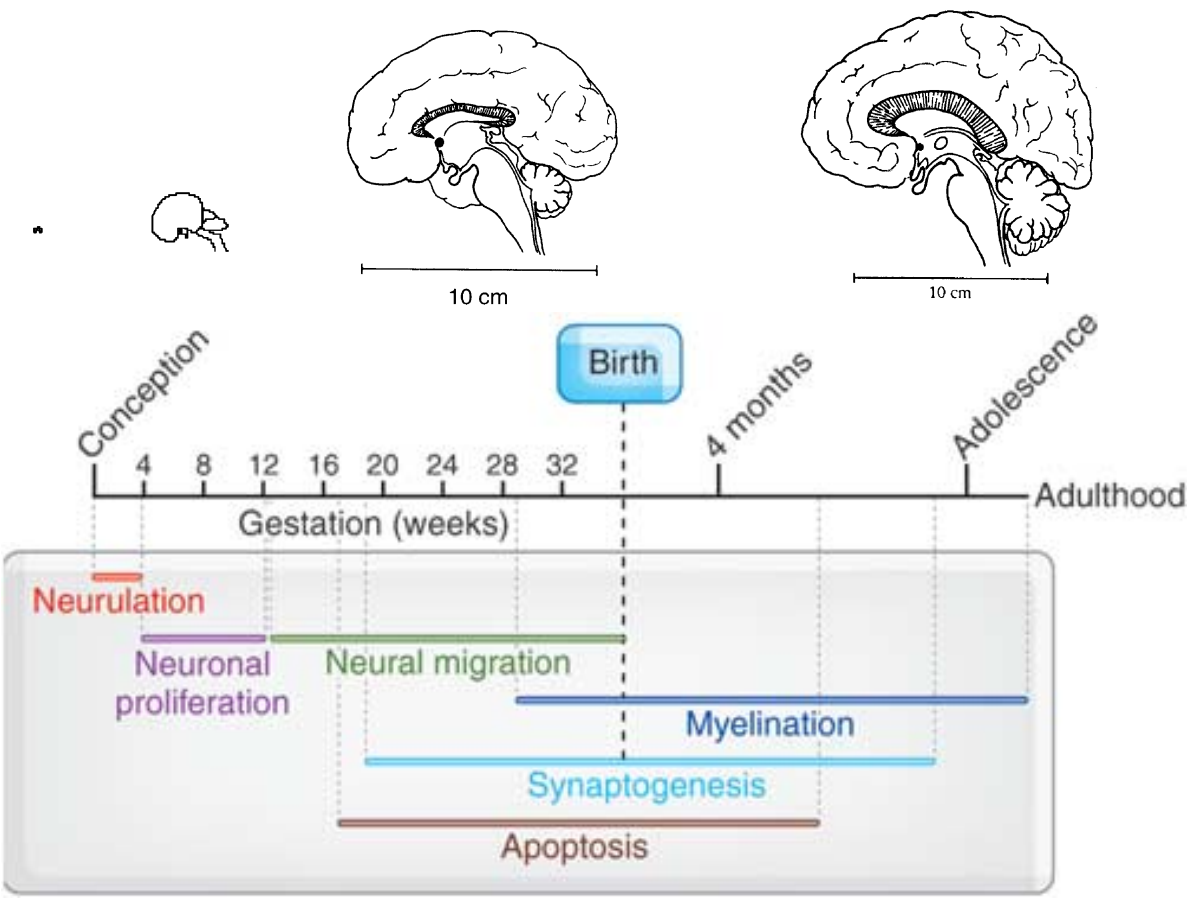
How epigenetics can blur the line between nature and nurture.

BY SIDDHARTHA MUKHERJEE

The author's mother (right) and her twin are a study in difference and identity.

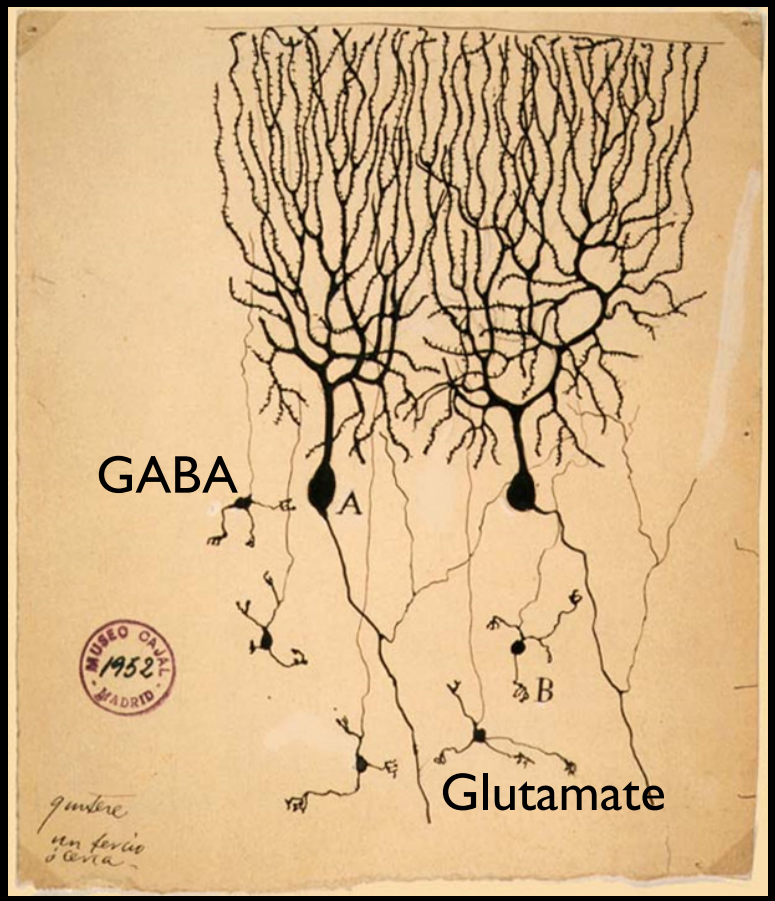
How do cells acquire, maintain and adjust their diverse characters?

Brain development:
A tightly orchestrated process



Tau and Peterson, B.S. *Neuropsychopharmac.* (2009).
S.L. Anderson. *Neurosci. and Behav. Rev.* (2003)

Puzzle:
Different cells, same genome...

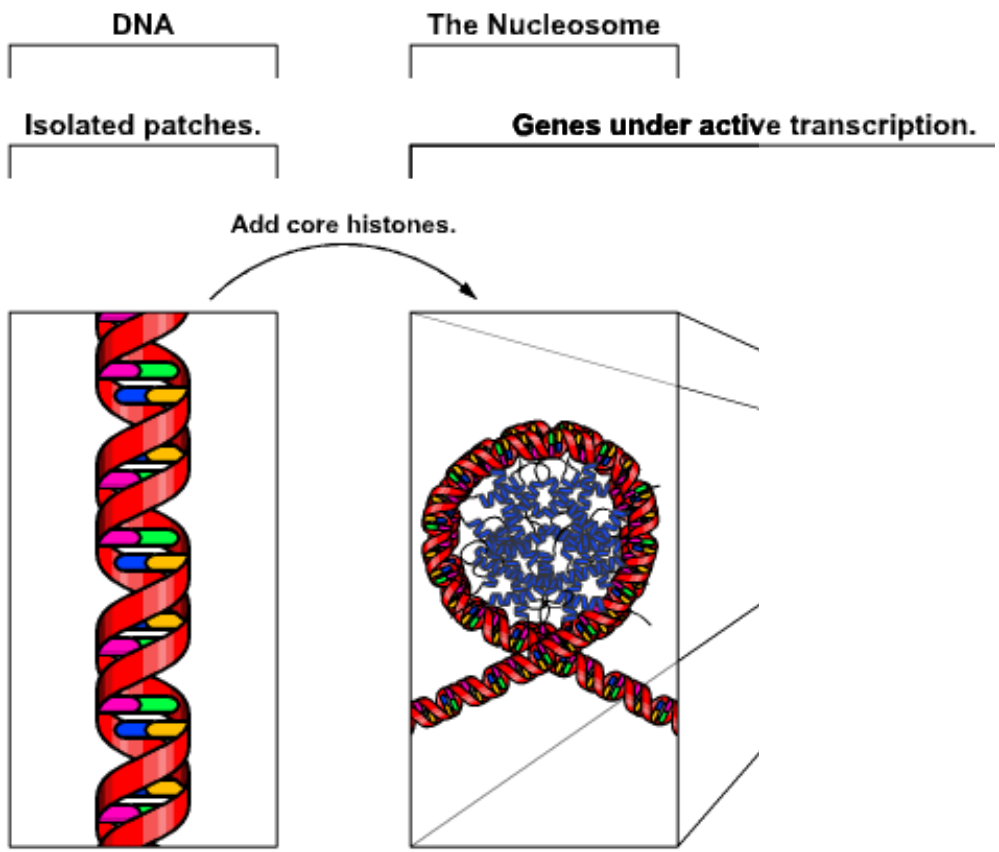


Drawing of pigeon cerebellar Purkinje and granule cells, Ramon y Cajal (1899)

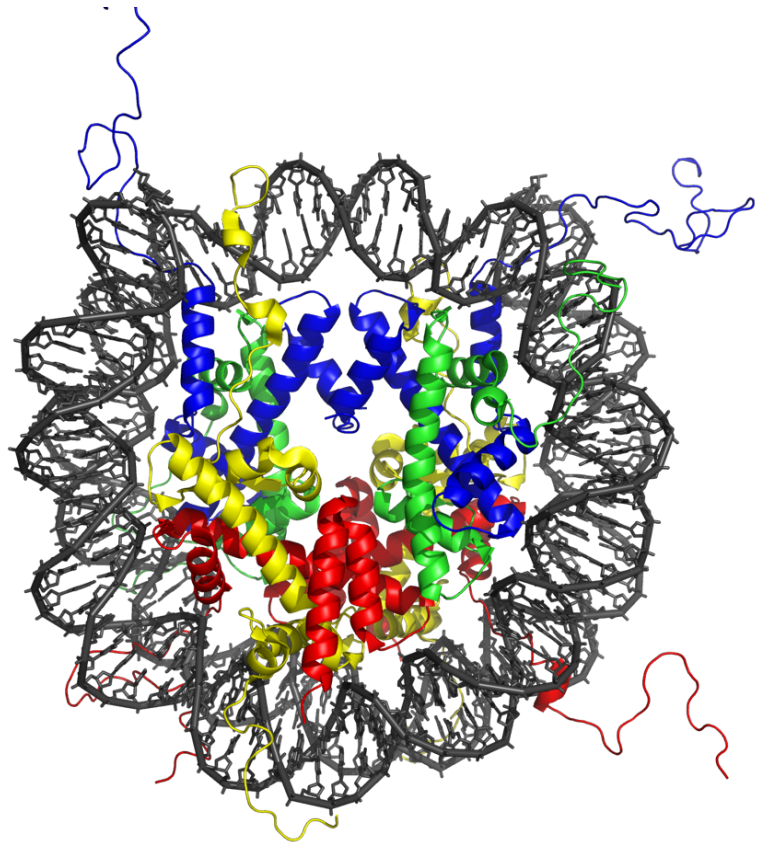
Same genome

Multiple mature phenotypes

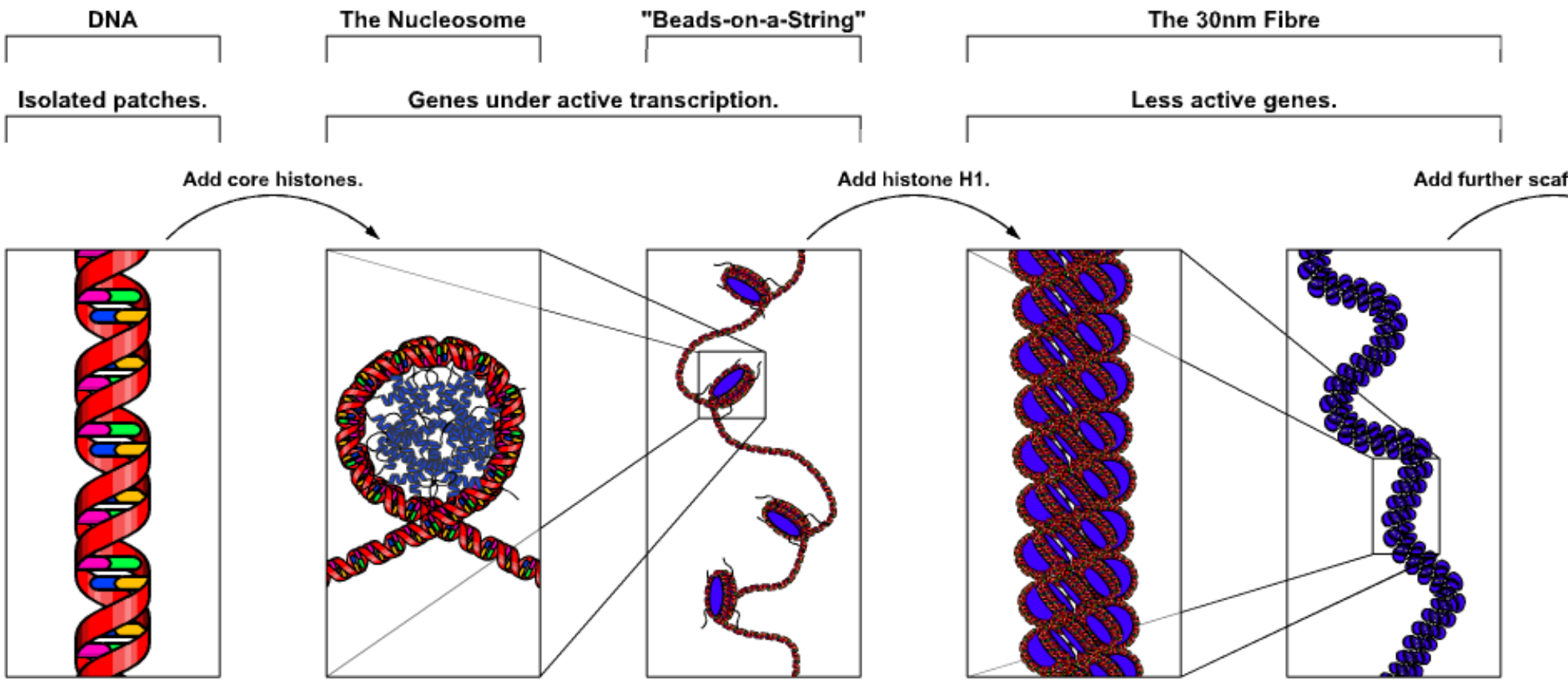
Levels of organization in the genome



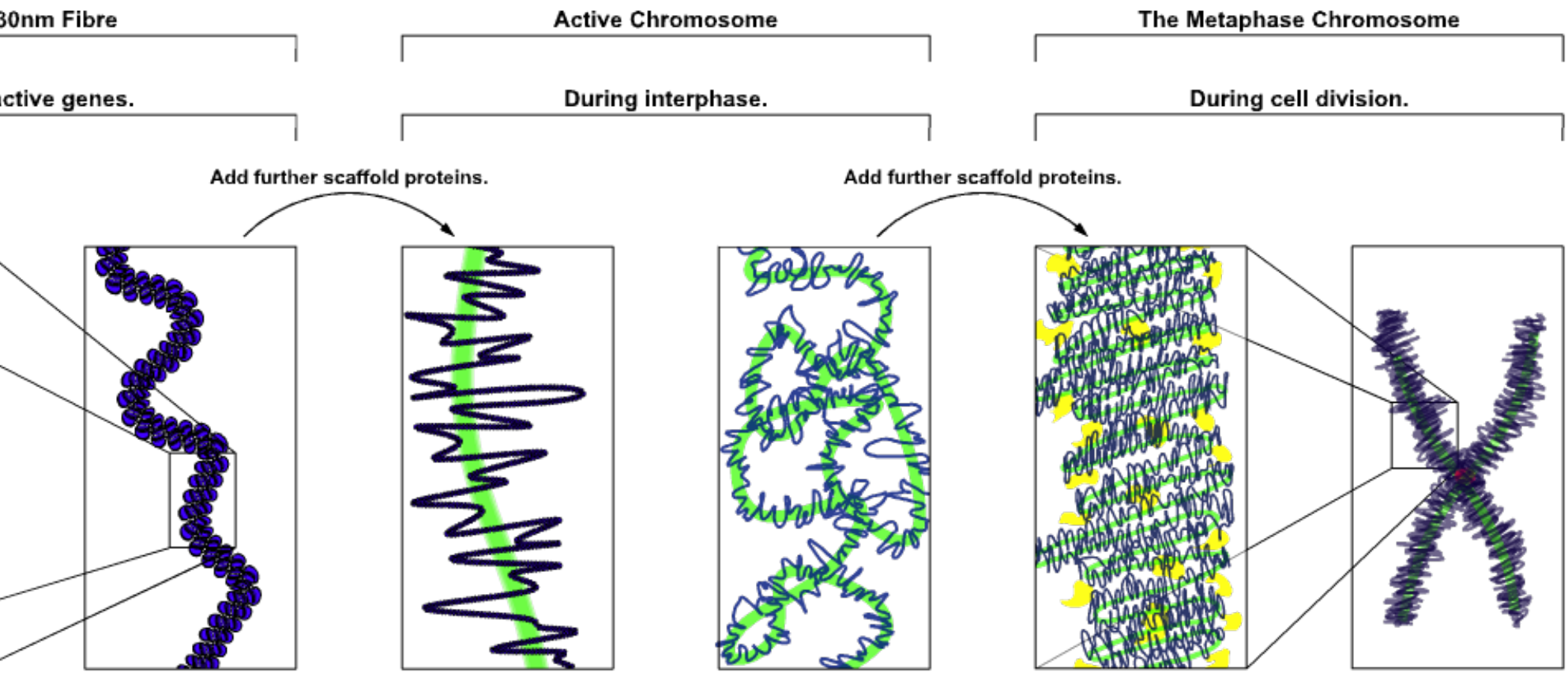
DNA wrapped around a histone particle



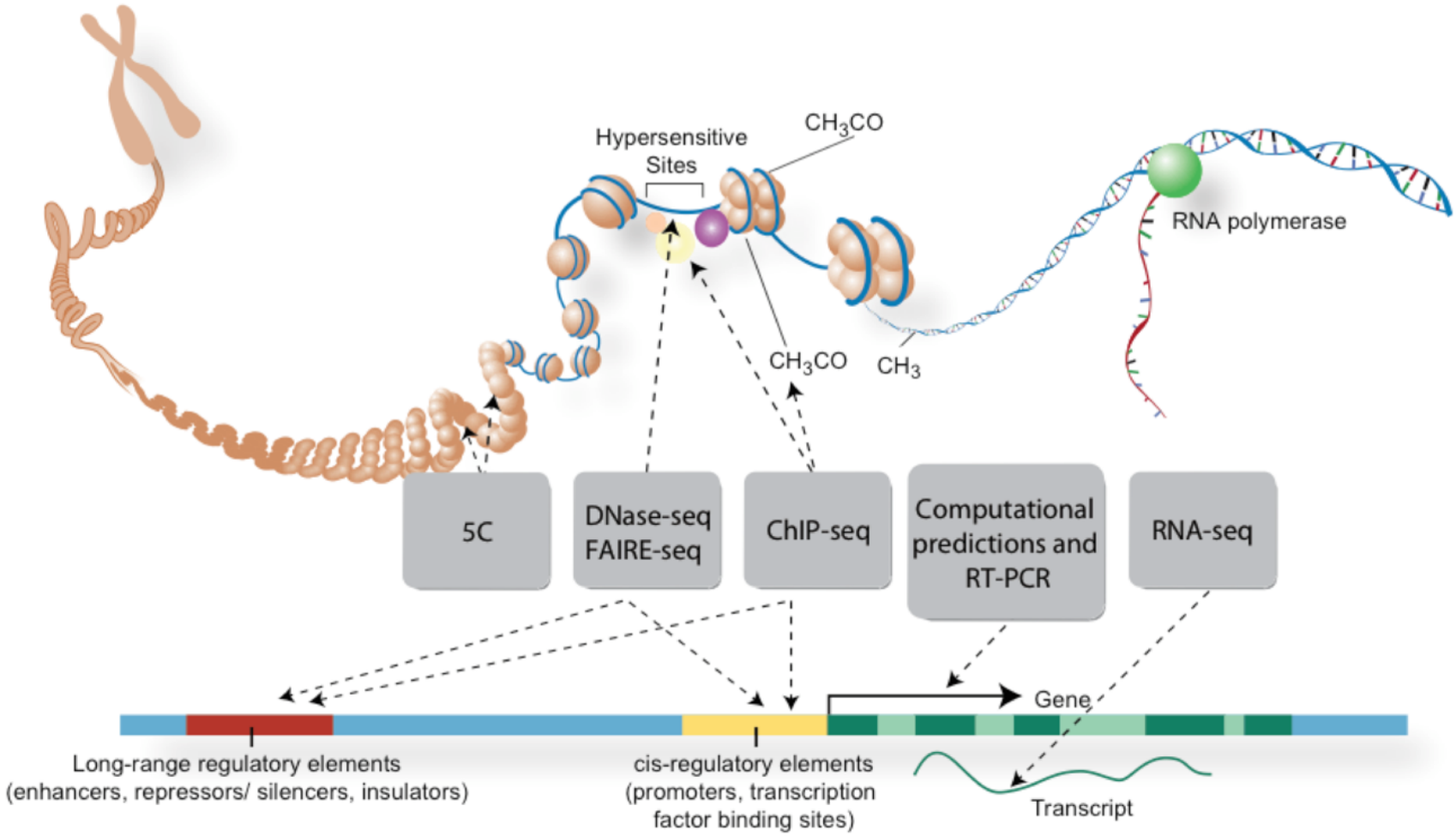
Levels of organization in the genome



Levels of organization in the genome



Modifications to DNA are epigenetic "punctuation marks"

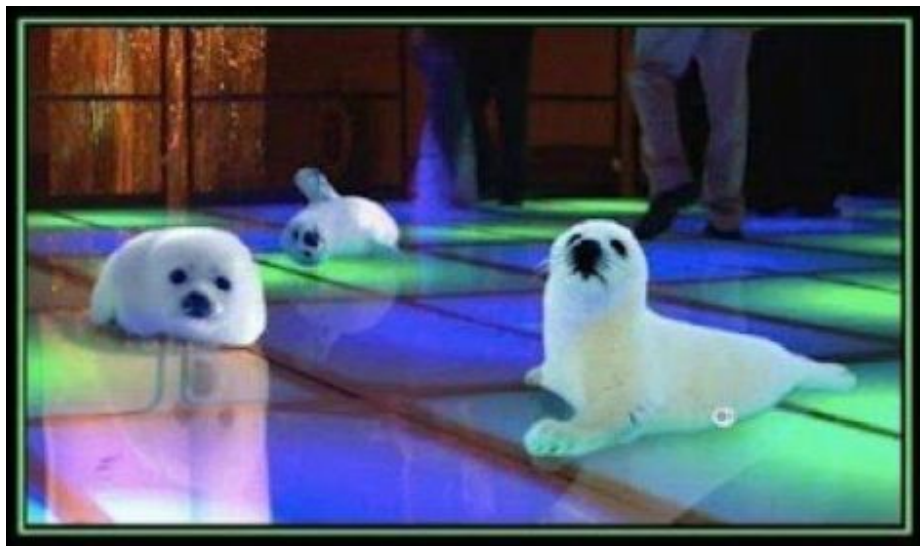


Encyclopedia of DNA Elements (ENCODE) (2012)

Epigenetics: Punctuation. Is. Key.

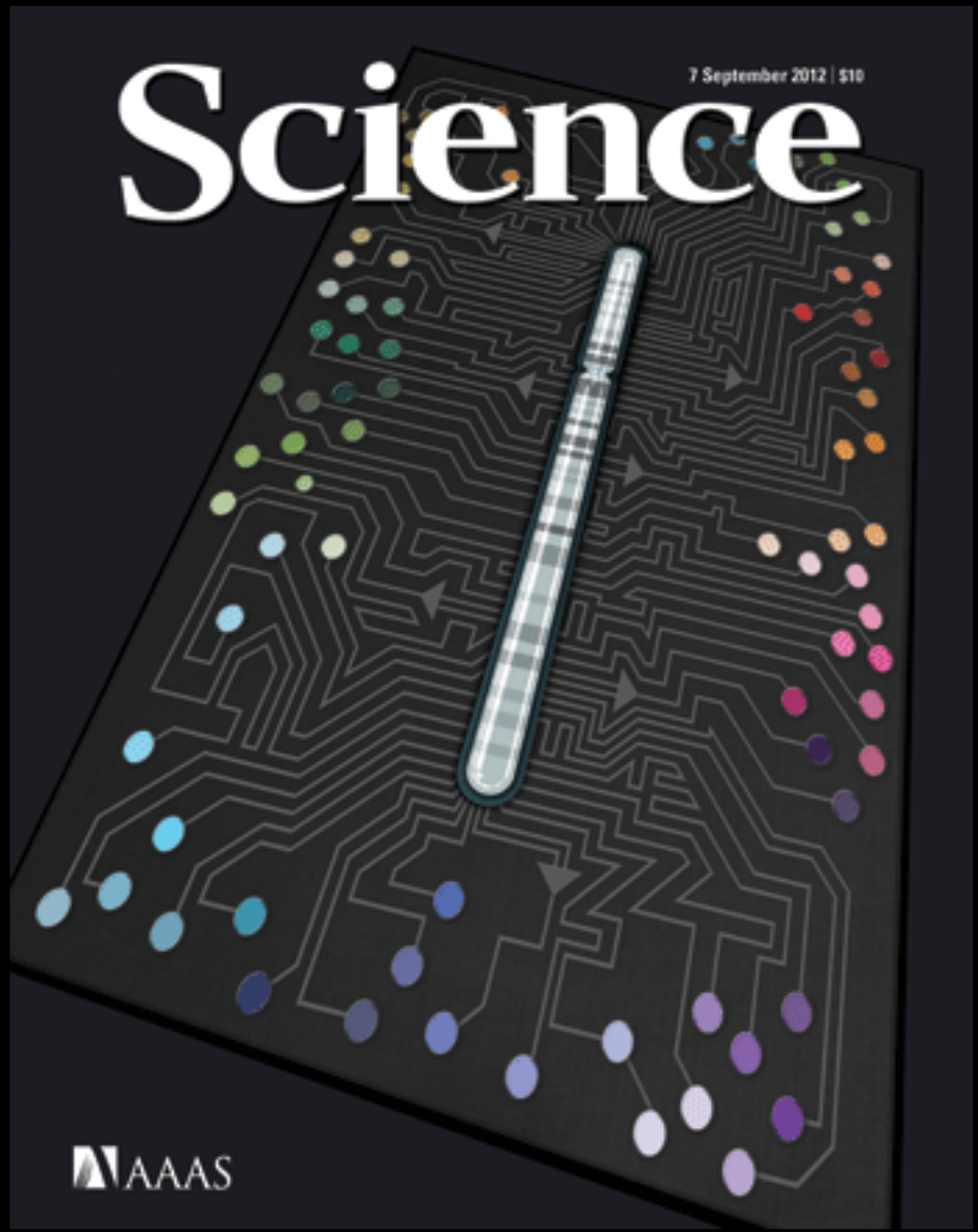
STOP CLUBBING BABY SEALS!!

STOP CLUBBING, BABY SEALS!

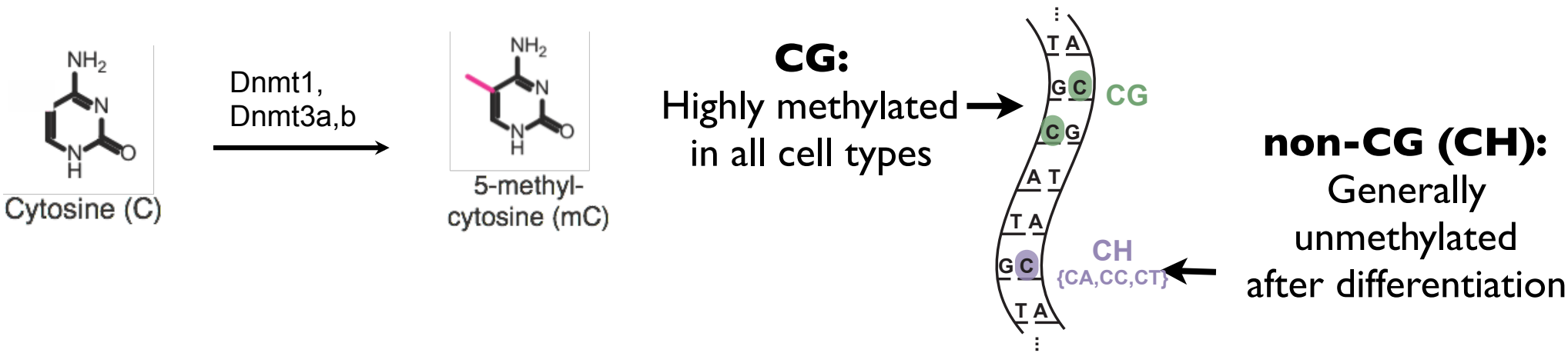


The genome as a computational network

ENCODE:
Encyclopedia of
DNA elements



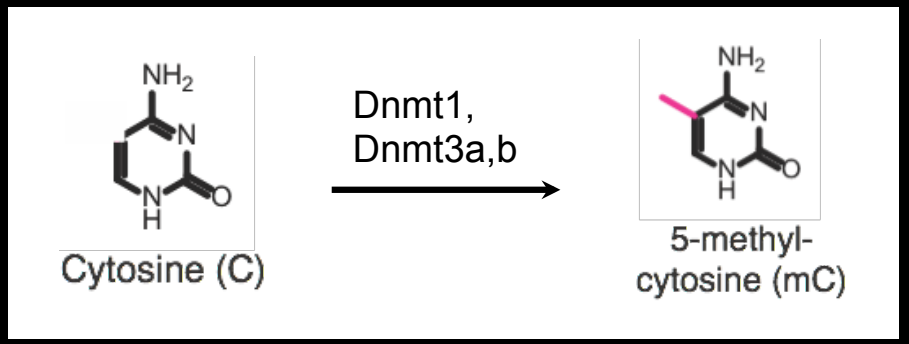
Cytosine DNA Methylation



- Covalent modification of genomic cytosine (mC)
- Key roles in imprinting, X-inactivation, transcription repression, cancer
- Stable and heritable
- Yet, reversible and potentially activity-dependent
- *Rett syndrome:* An autism-spectrum disorder caused by Methyl-C Binding Protein (MECP2) loss of function

Epigenetic regulation in insects

Queen and female worker bees



Royal jelly

Plants

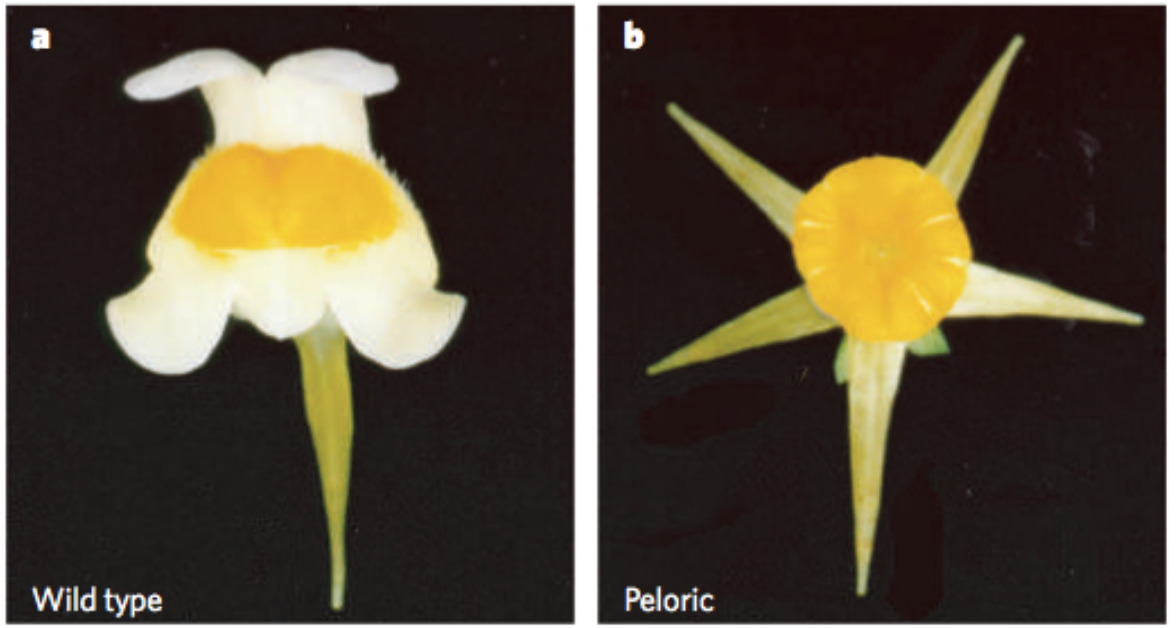


Figure 1 | Frontal view of a wild-type toadflax flower and a peloric epimutant. a, The wild-type flower is dorsoventrally asymmetrical. **b,** By contrast, the peloric flower is radially symmetrical with all petals resembling the ventral petal of the wild-type flower. (Image reprinted, with permission, from ref. 4.)

Epimutation:

Two forms of the toadflax plant with identical genotype but different inherited DNA methylation patterns

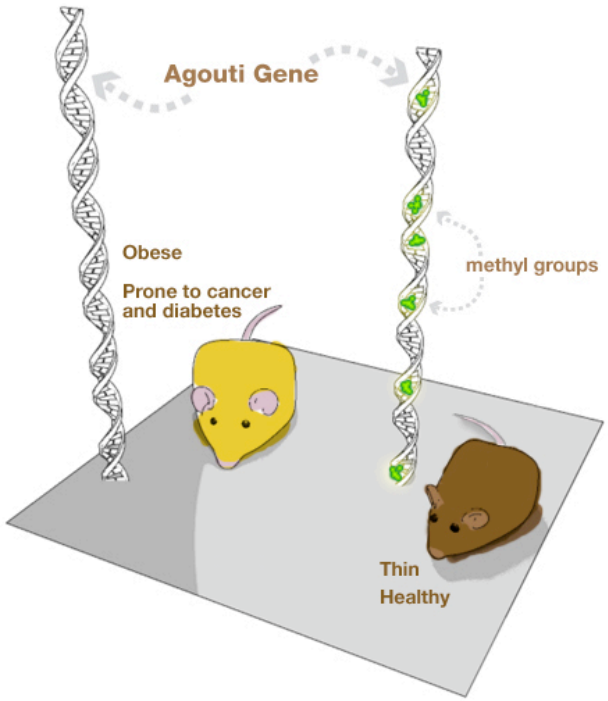
Bird, A. (2007). Perceptions of epigenetics. *Nature*, *447*(7143), 396–398. doi:10.1038/nature05913

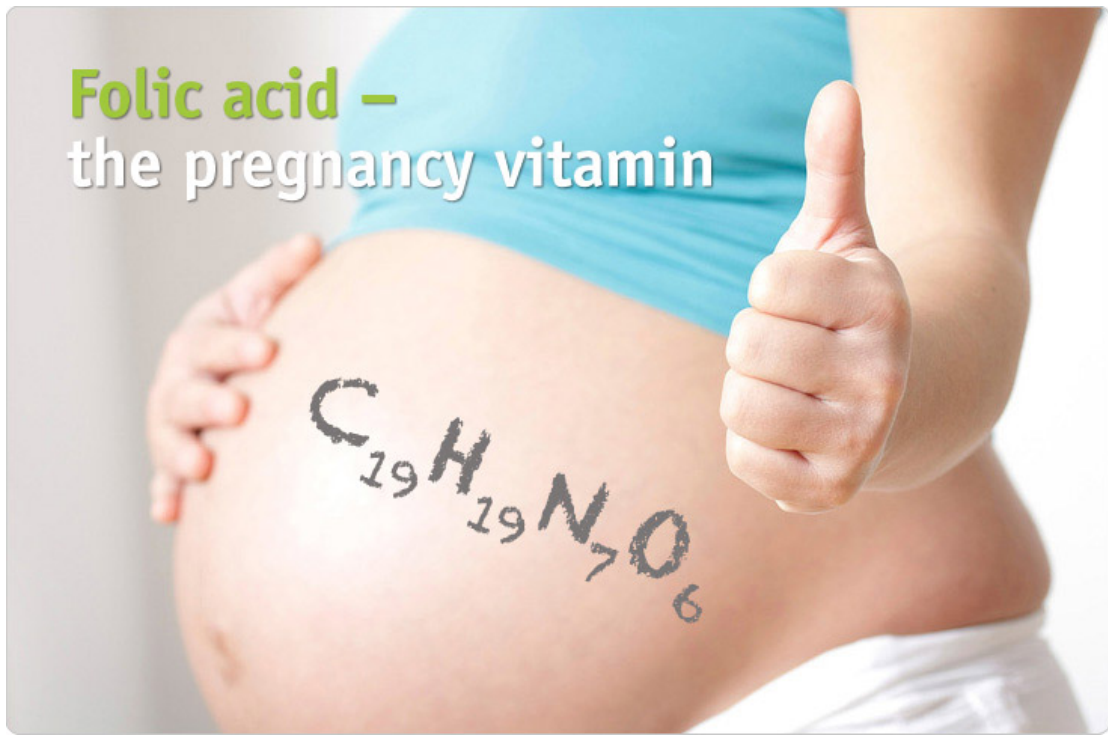
Mammals: You are what you eat (so eat your vitamins!)

Nutrients supporting healthy methylation:

- Folic acid
- B-vitamins
- SAM (S-adenosyl methionine)

Especially important for pregnant mothers/infants





Maternal care induces life-long changes in DNA methylation and stress resilience in offspring

Low quality maternal care

↑ Stress behavior of offspring

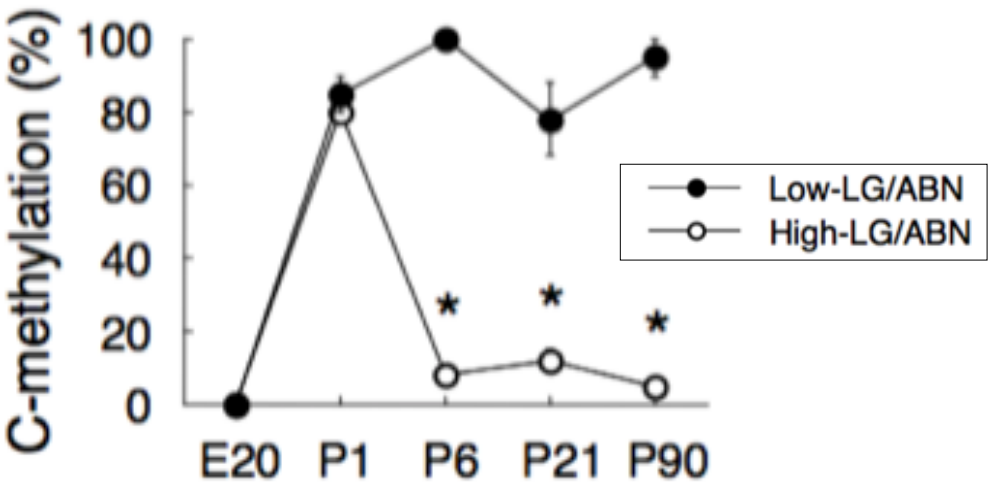
↓ Glucocorticoid receptor (GR) expression in brain

High quality maternal care

↓ Stress behavior of offspring

↑ Glucocorticoid receptor (GR) expression in brain

GR promoter is hypermethylated in hippocampus in low-quality group [Weaver, …, Szyf, Meaney (2004)]



Maternal Care

↓

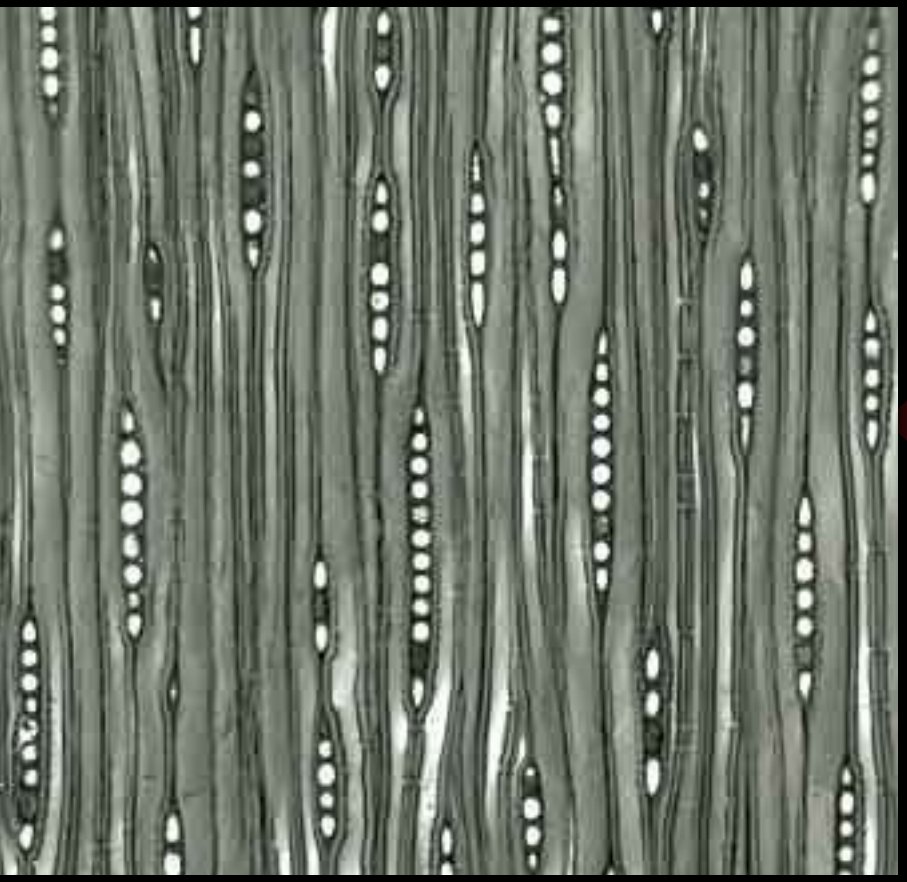
Offspring DNA methylation

↓

Offspring GR expression

↓

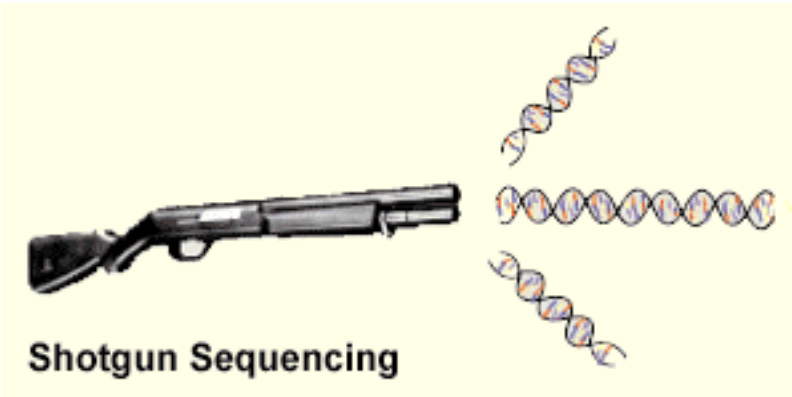
Offspring stress behavior

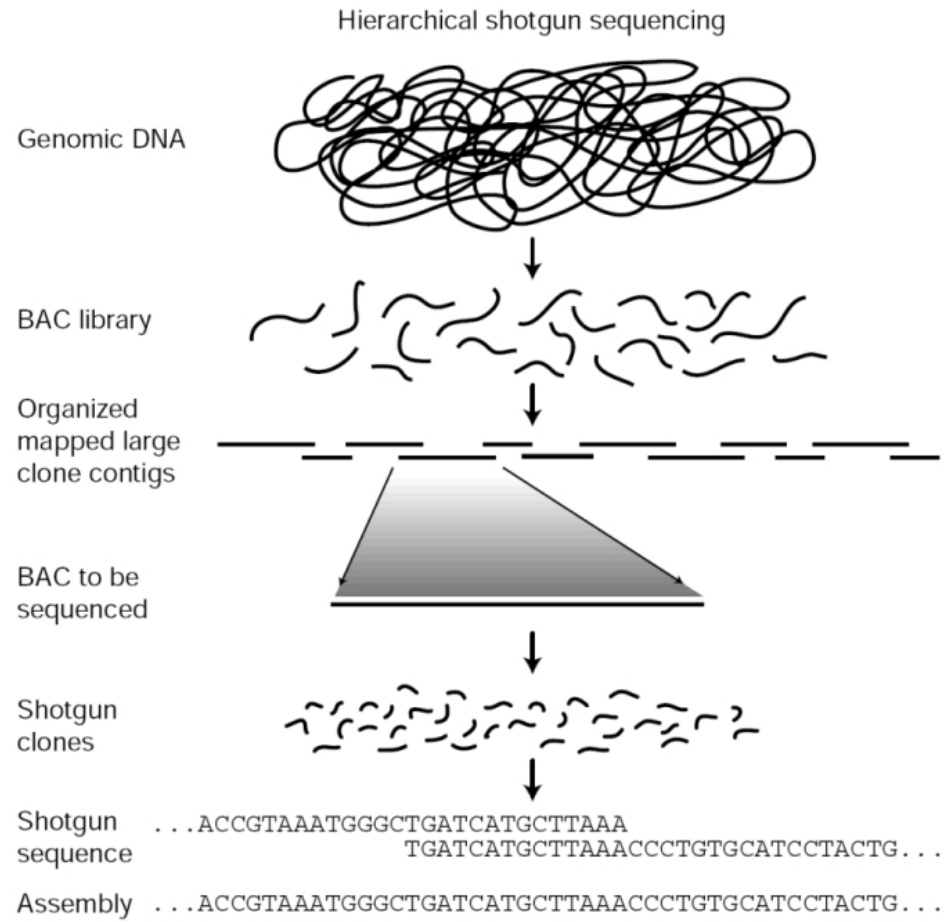


~10^6 x

Testing the cognitive role of epigenetic modifications requires **genome-scale, base-resolution neuronal epigenome profiling**

Studying gene networks: Shotgun sequencing



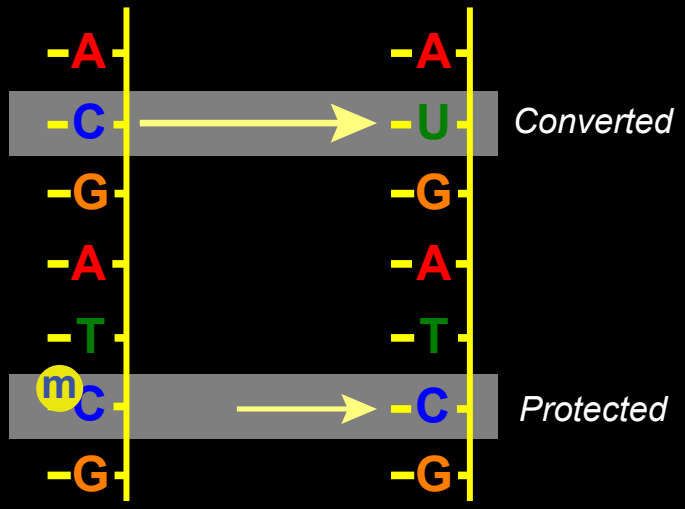


Shotgun bisulfite sequencing measures the DNA methylation landscape

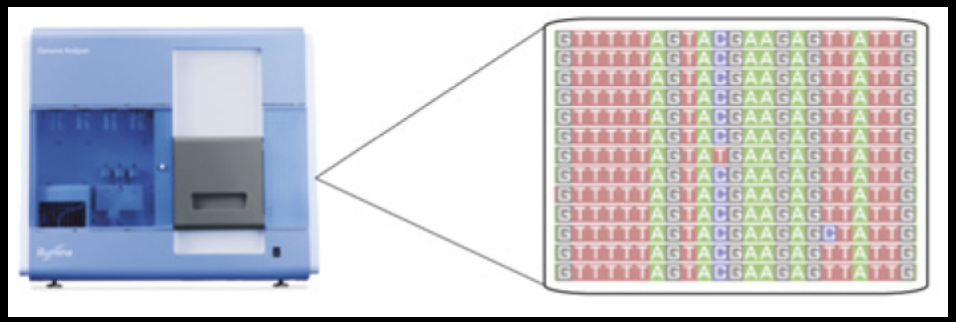
Genomic DNA

Random fragmentation

Bisulfite conversion of unmethylated C



Deep sequencing



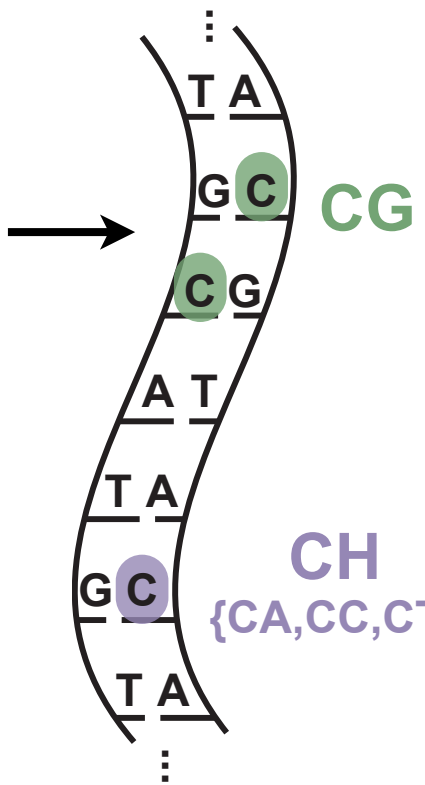
Computational analysis/statistics

$$c(\tau) = \sum_i m_i m_{i+\tau}$$

R. Lister and J. Ecker, *Genome Research* (2009)

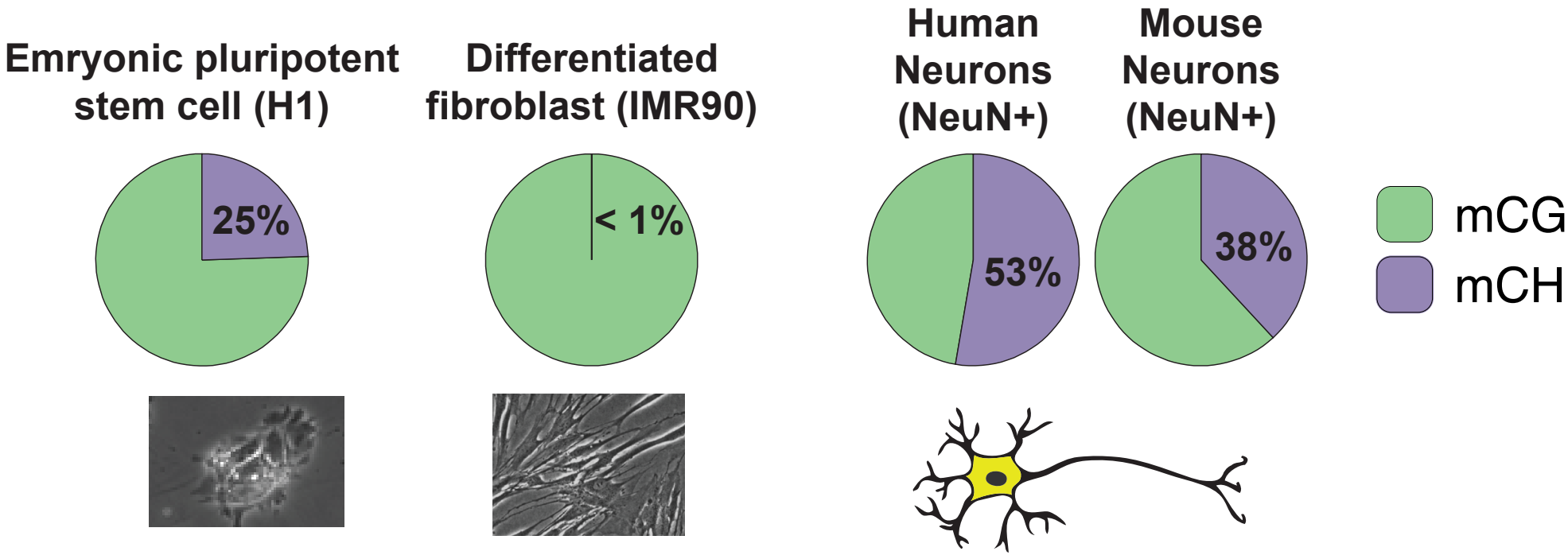
Two DNA sequence contexts for methylation

CG:
Highly methylated
in all cell types



non-CG (CH):
Generally
unmethylated
after differentiation

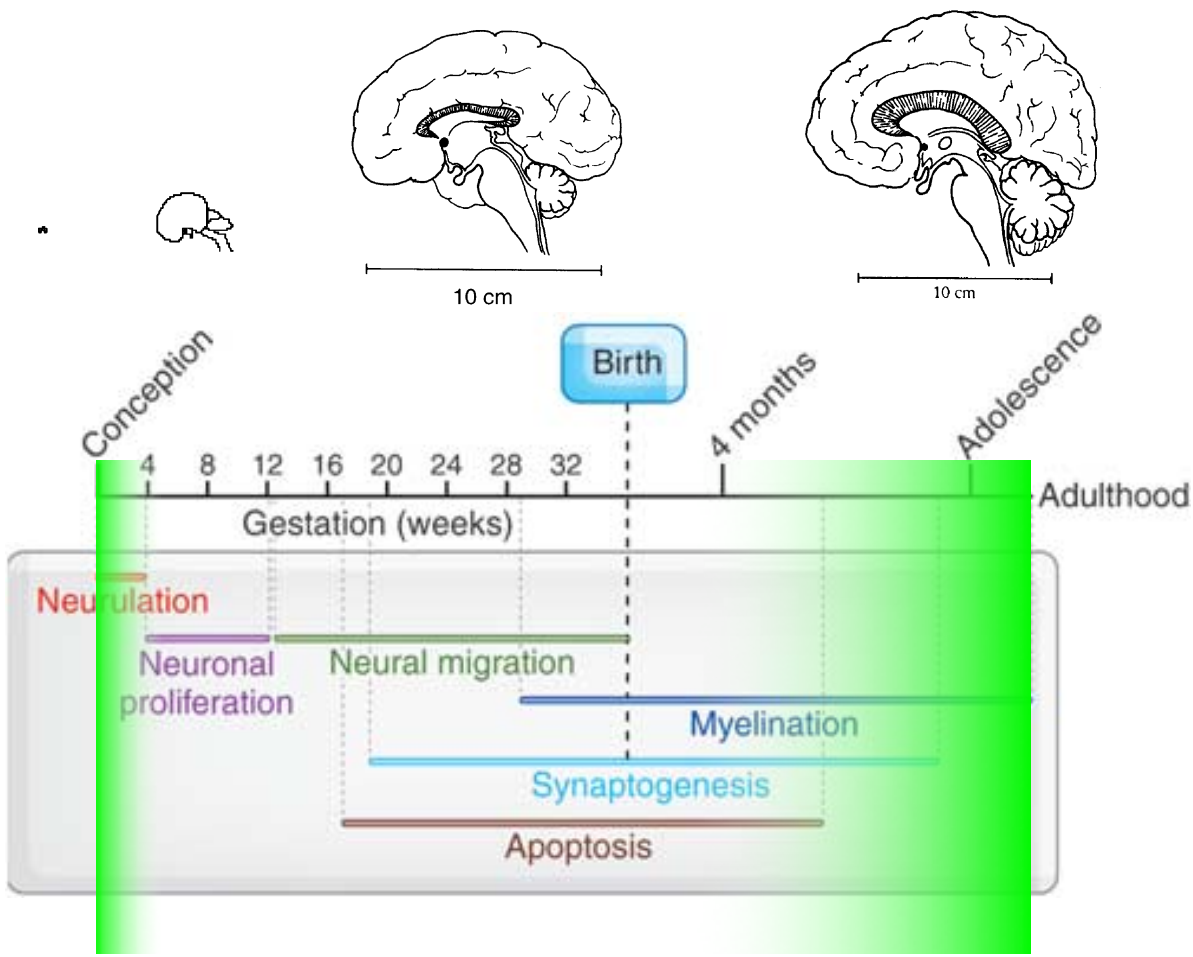
A surprise: Substantial non-CG methylation in neurons

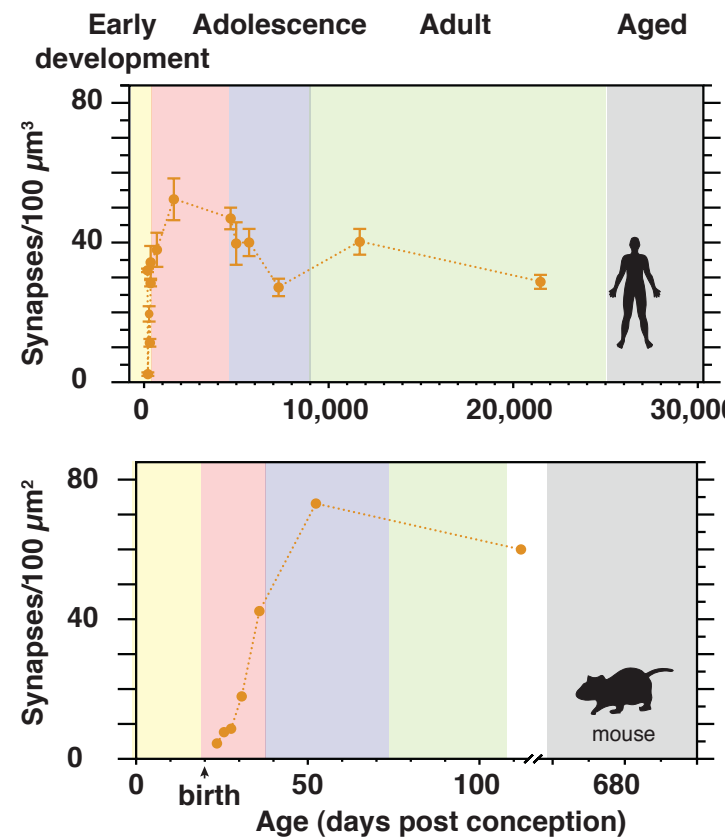


R. Lister*, E.A, Mukamel* *et al.*. *Science* (2013)

See also: Xie *et al.*, *Cell* (2012); Zeng *et al.*, *Am. J. Hum. Gen.* (2012)

How does methylation accumulate during brain development?



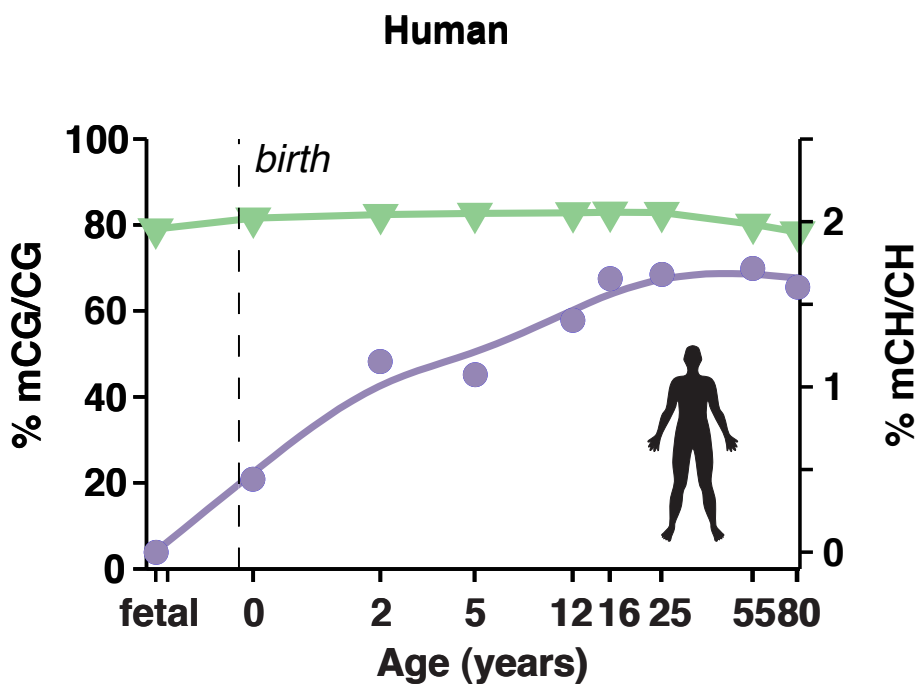


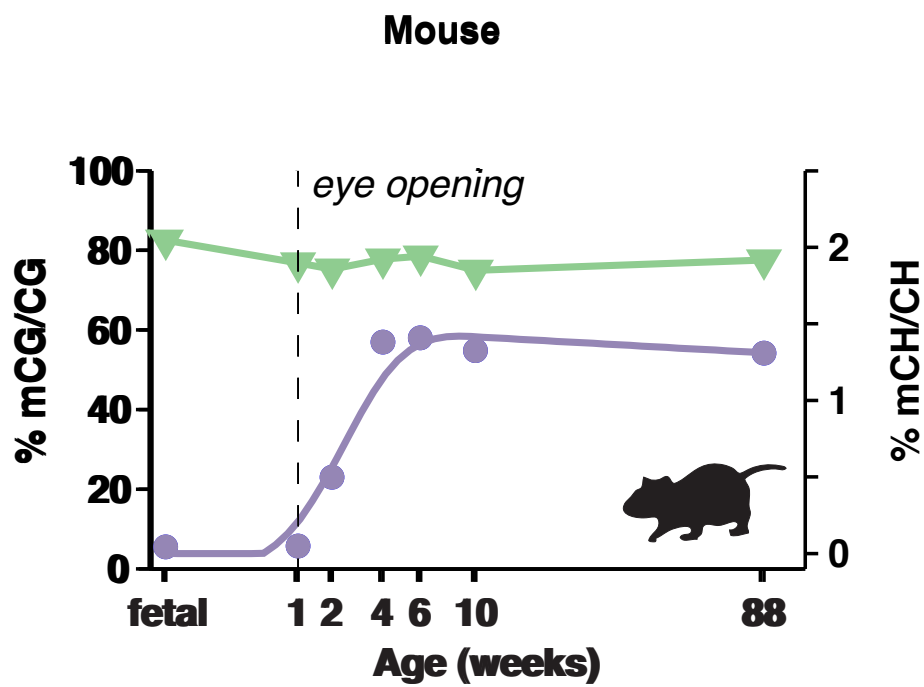
mCH **?** **mCH**

Tau and Peterson, B.S. *Neuropsychopharmac.* (2009).
S.L. Anderson. *Neurosci. and Behav. Rev.* (2003)

DeFelipe *et al.* (1997);
Huttenlocher et al., (1997)

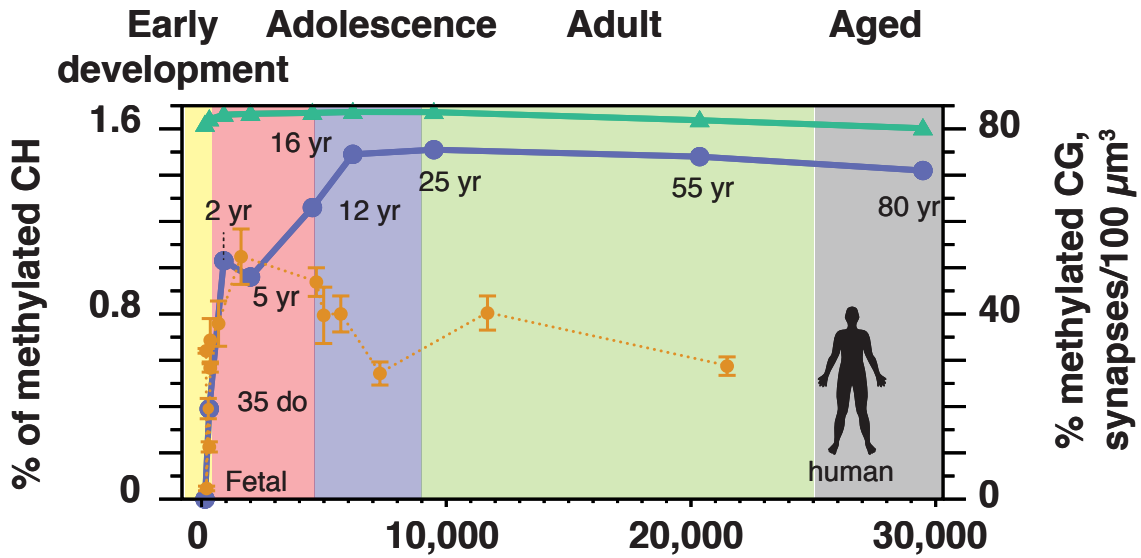
Non-CG methylation accumulates throughout childhood and adolescence





R. Lister*, E.A, Mukamel* *et al.*. *Science* (2013)

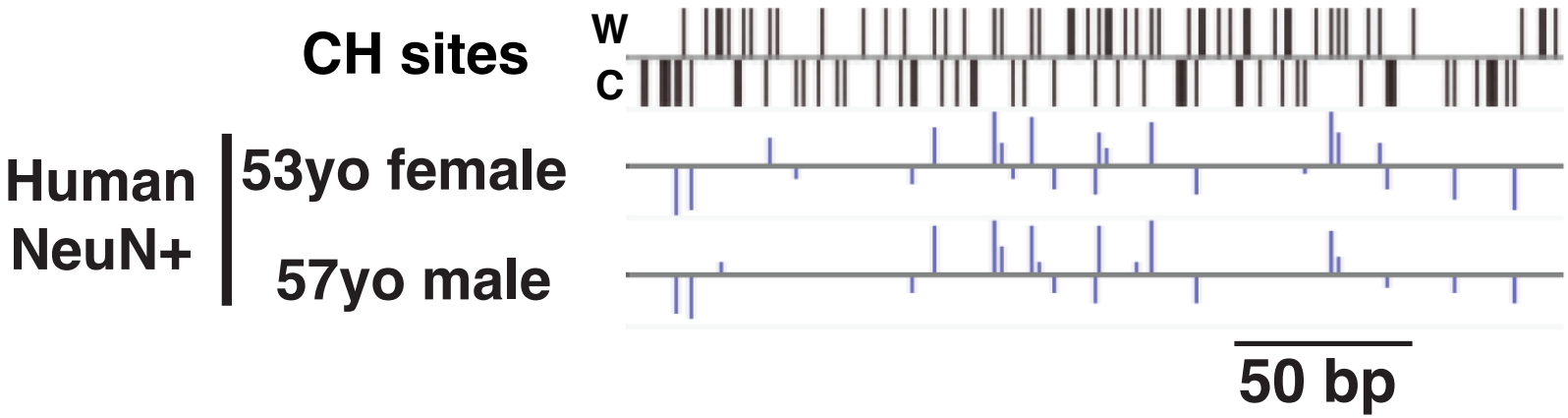
Non-CG methylation increases during years 0-16, coinciding with synaptogenesis and pruning



*DeFelipe *et al.* (1997); Huttenlocher et al., (1997)

R. Lister*, E.A, Mukamel* *et al.*. *Science* (2013)

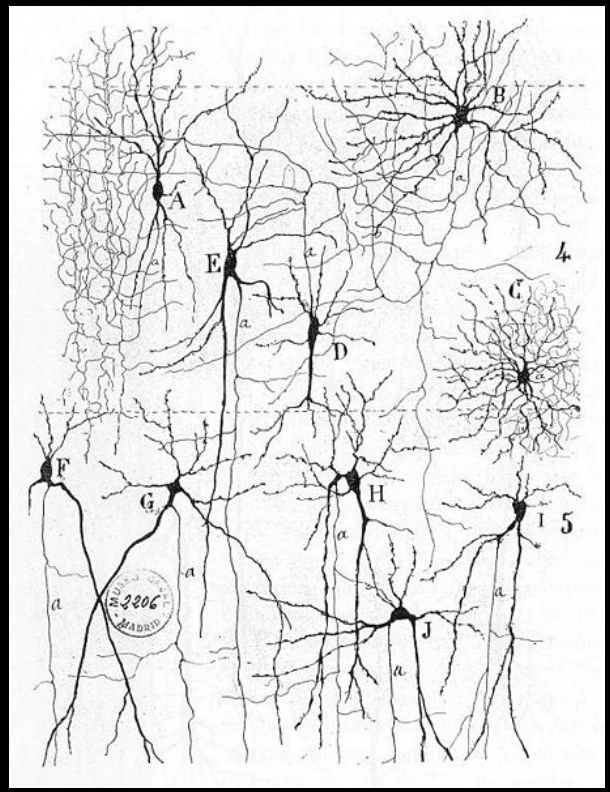
Methylation patterns are strongly conserved between individuals



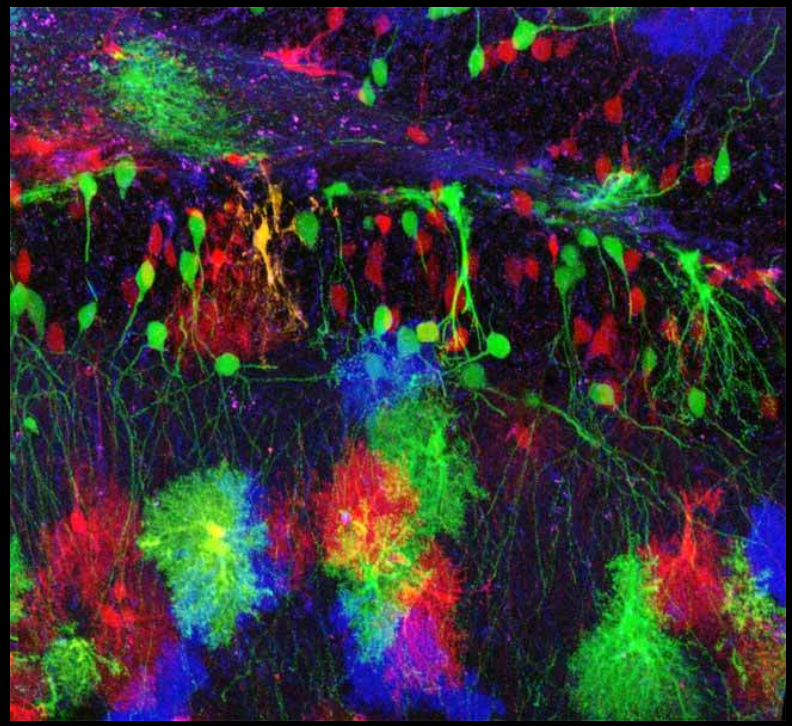
Conservation suggests there could be a biological function

However, there is no causal evidence yet (stay tuned)

Does DNA methylation contribute to brain cell diversity?

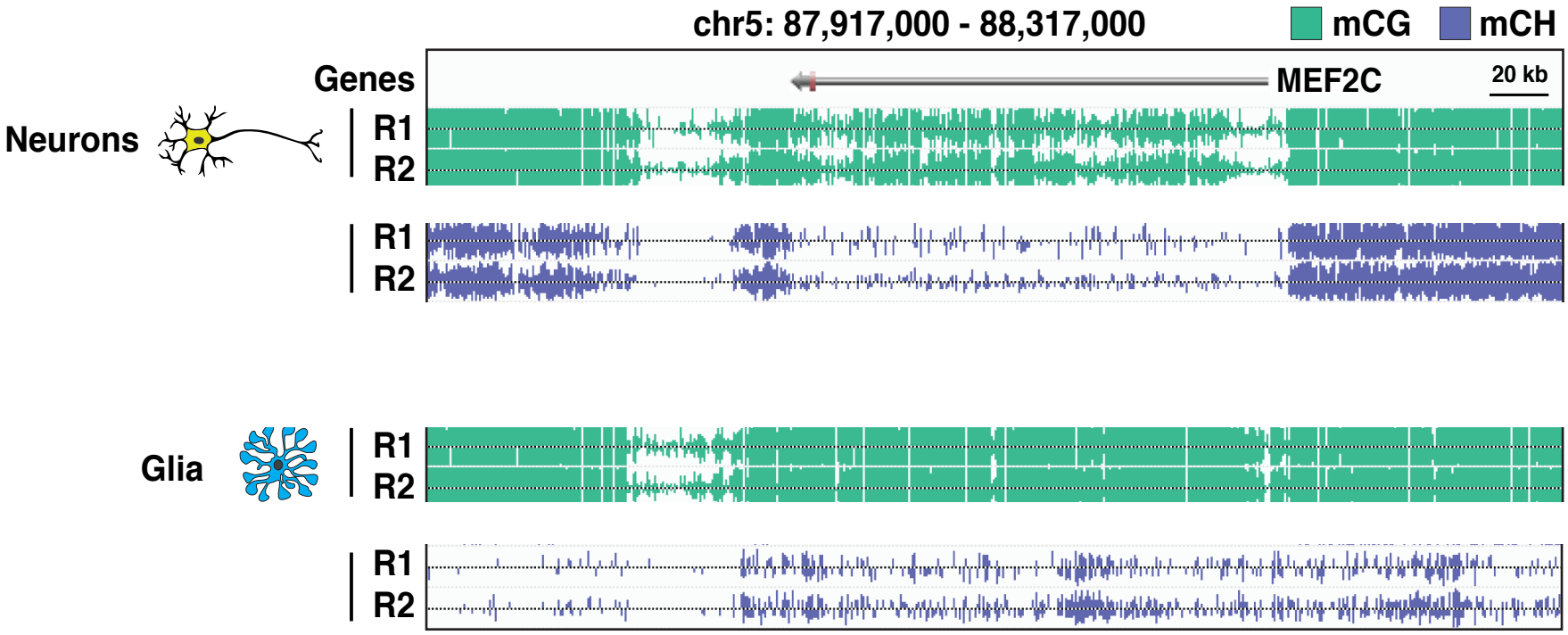


Drawing of auditory cortical neurons, Ramón y Cajal (1899)



Fluorescently labeled neurons and glia,
Livet, Sanes, and Lichtman (2007)

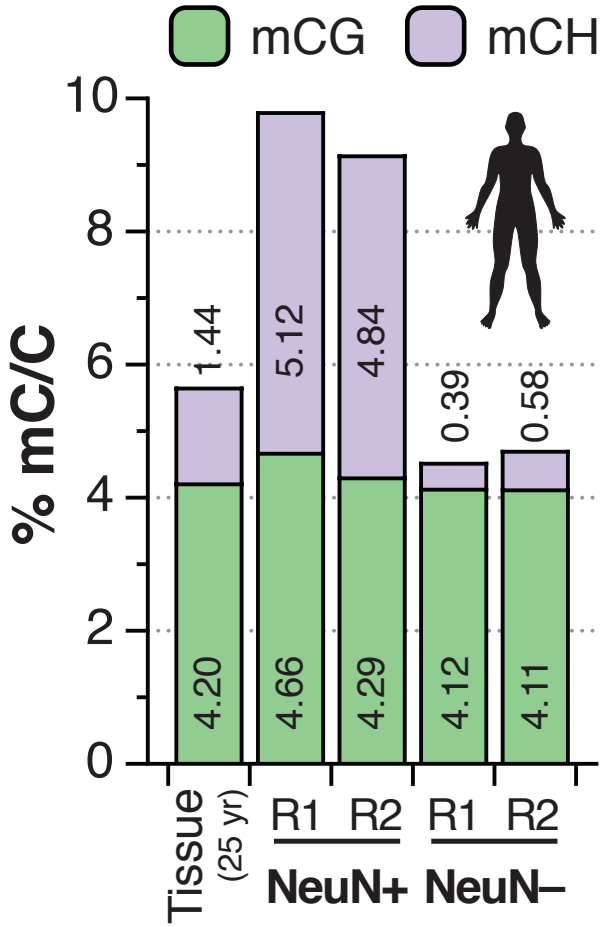
Cell types have unique DNA methylation fingerprints

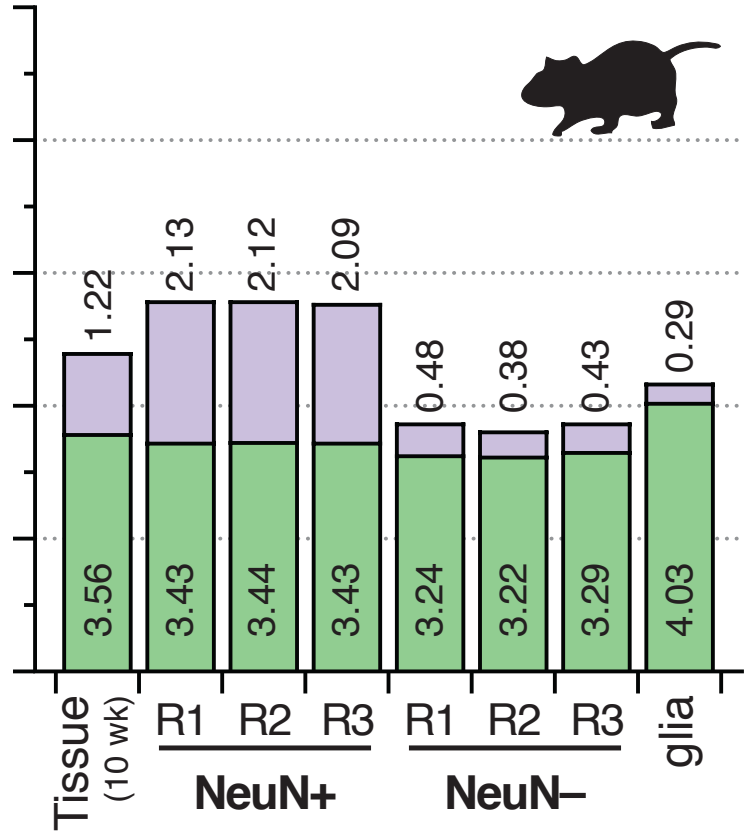


Transcription factor MEF2C:

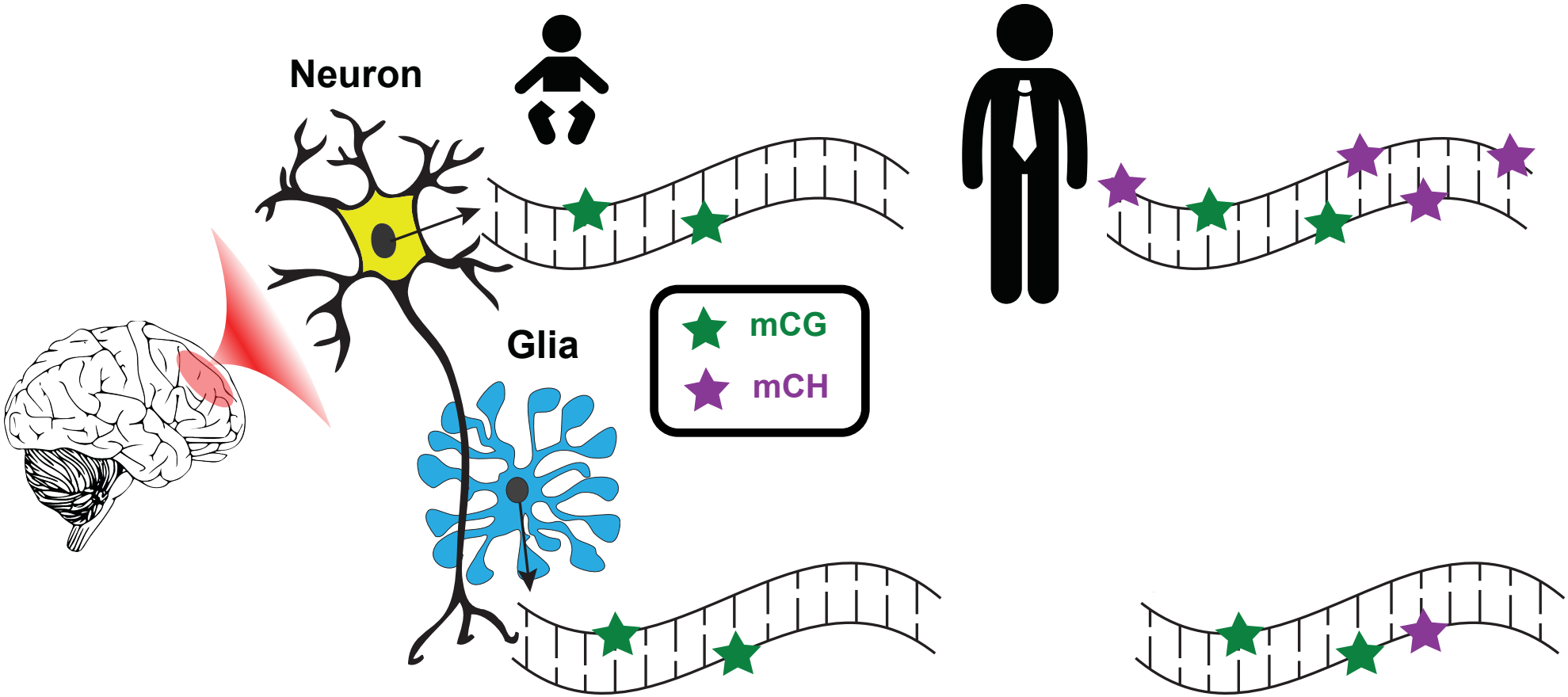
- Implicated in neurogenesis and cortical development
- Hypermethylated (*i.e.*, repressed) in glia

mCH is a characteristic feature of neurons, not astrocytes

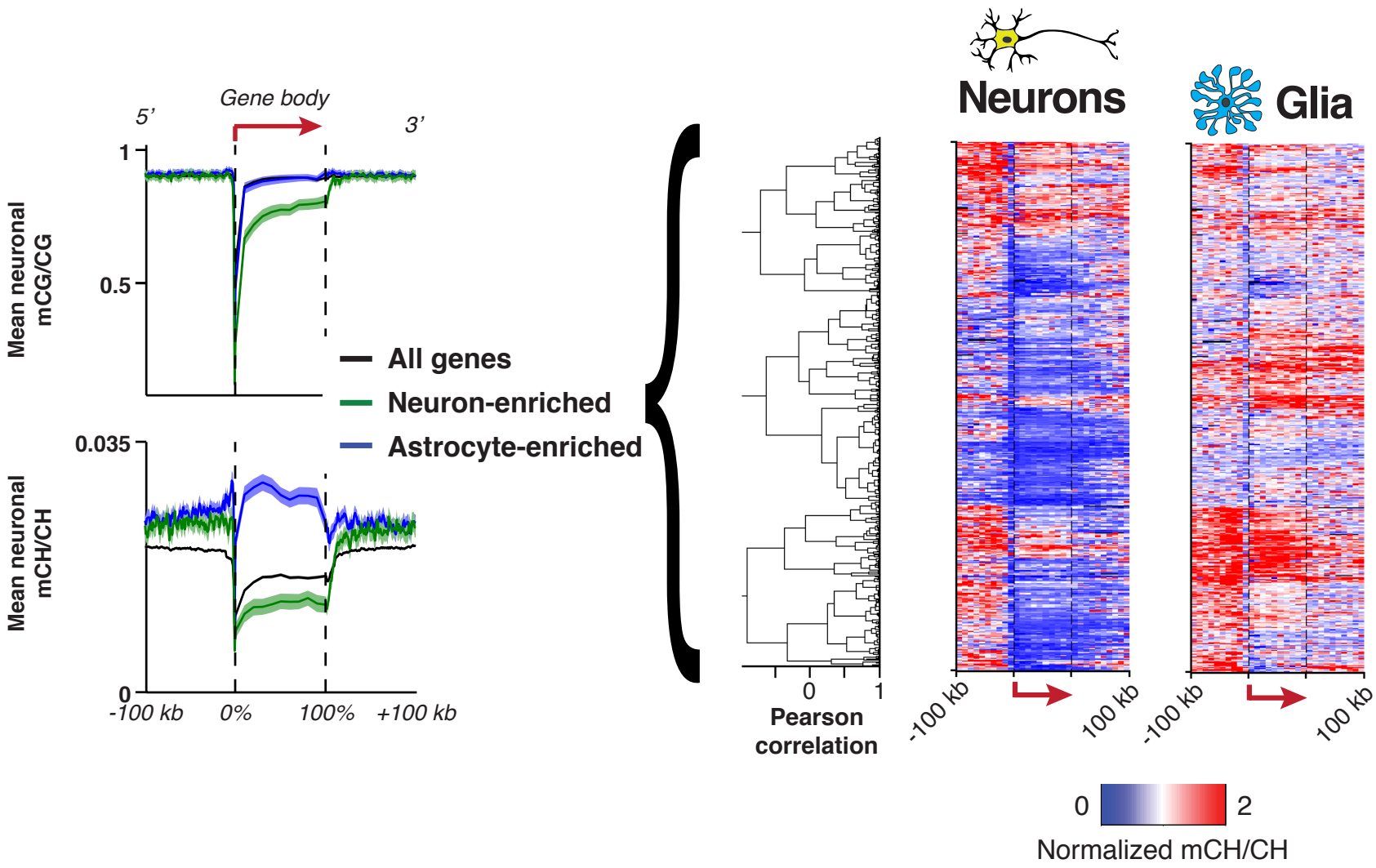




Non-CG DNA methylation is a specific feature of mature neurons

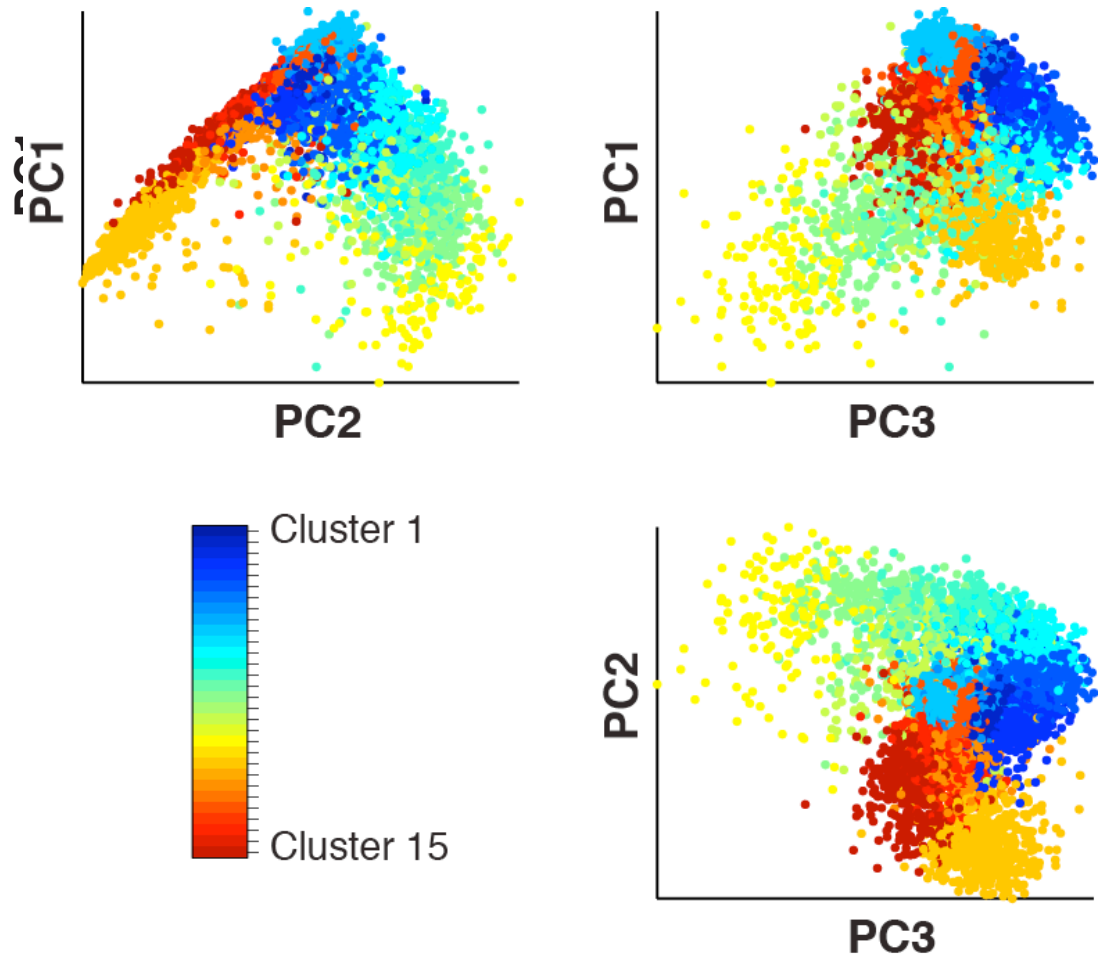


Identifying gene methylation patterns is a "Big Data" challenge



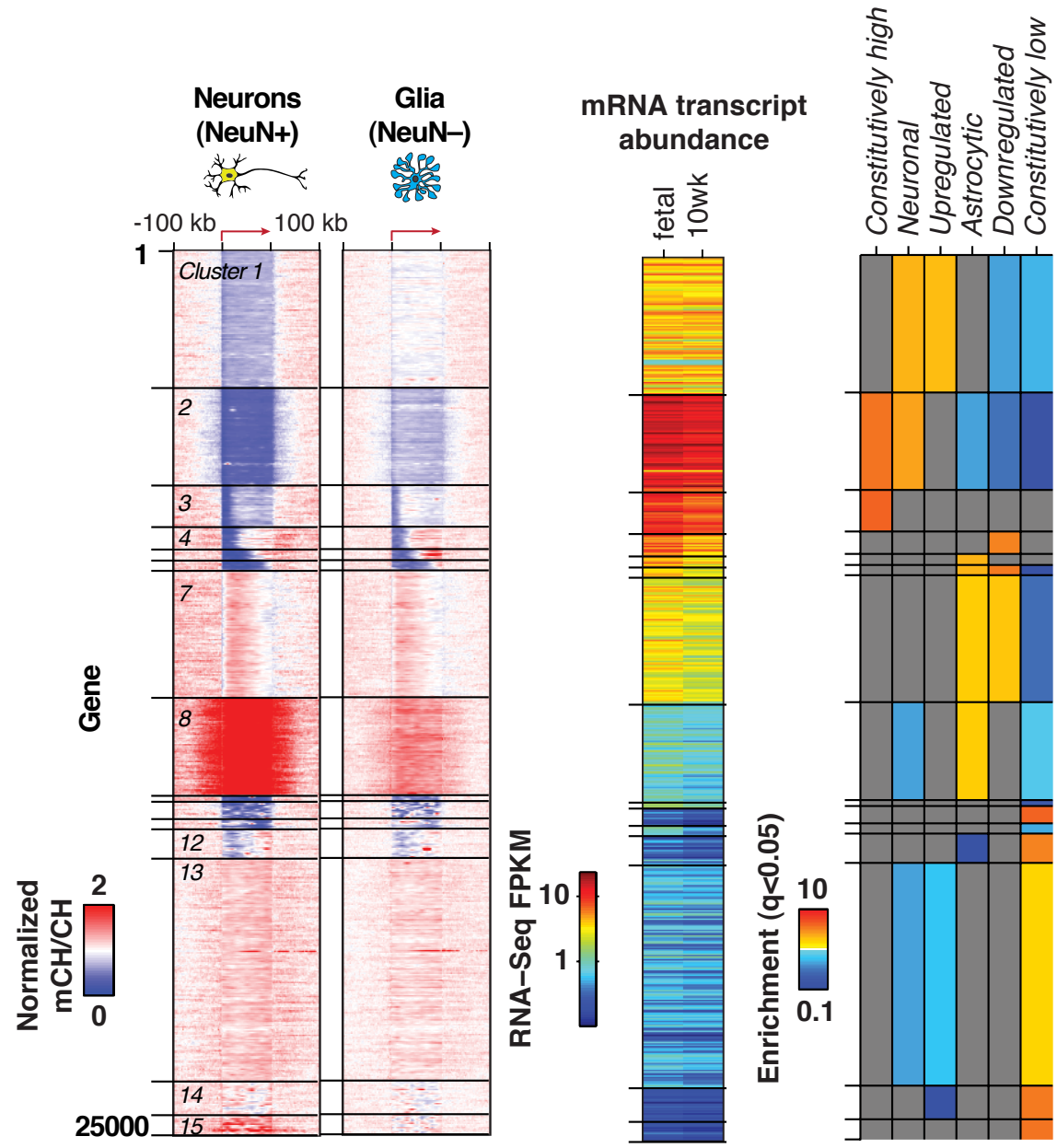
Lister*, Mukamel*, *et al.* (2013)

Unbiased clustering of methylation profiles identifies distinct gene sets

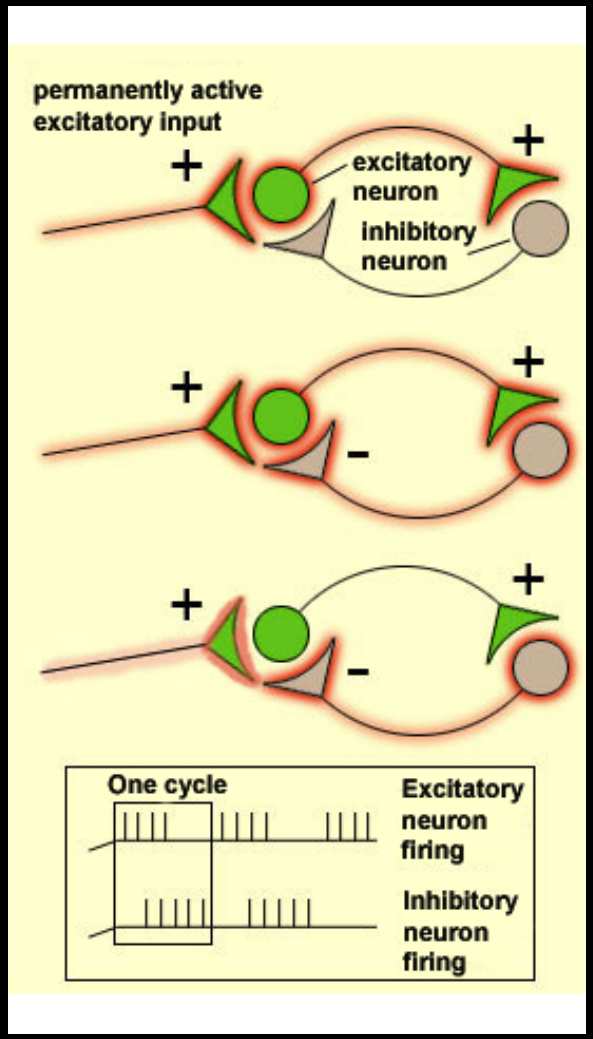


Principal component (PC) analysis of genome-wide methylation patterns

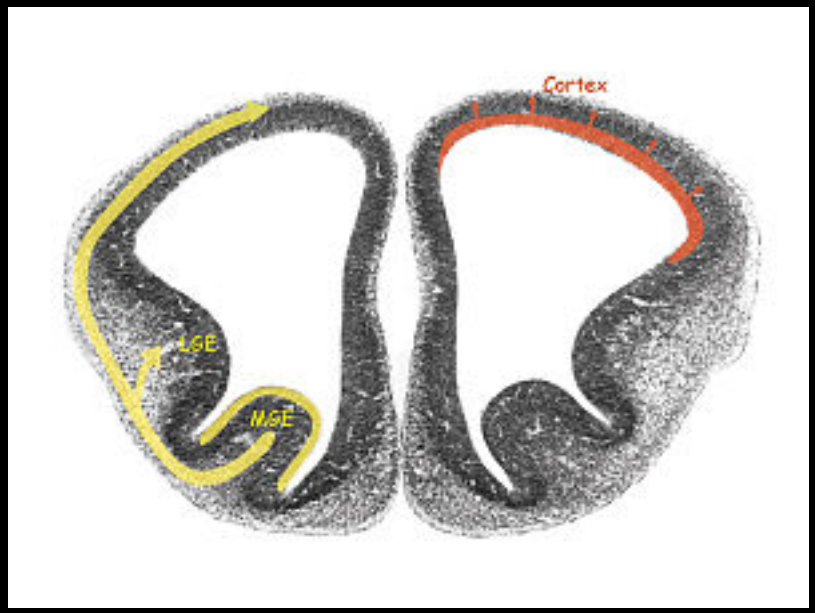
Unbiased clustering of methylation profiles identifies distinct gene sets

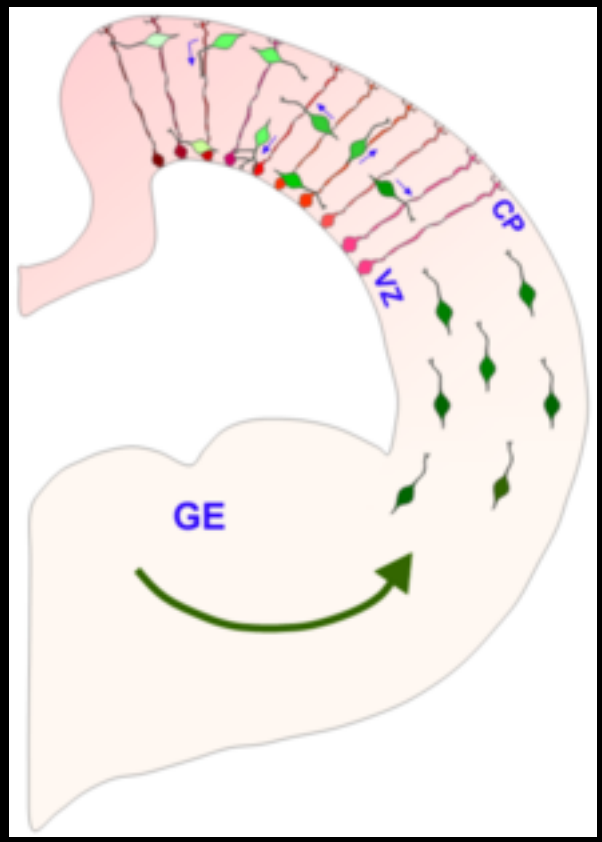


Sub-types of neurons:
Excitatory and inhibitory cells create balance



Excitatory and inhibitory neurons: Natives and immigrants

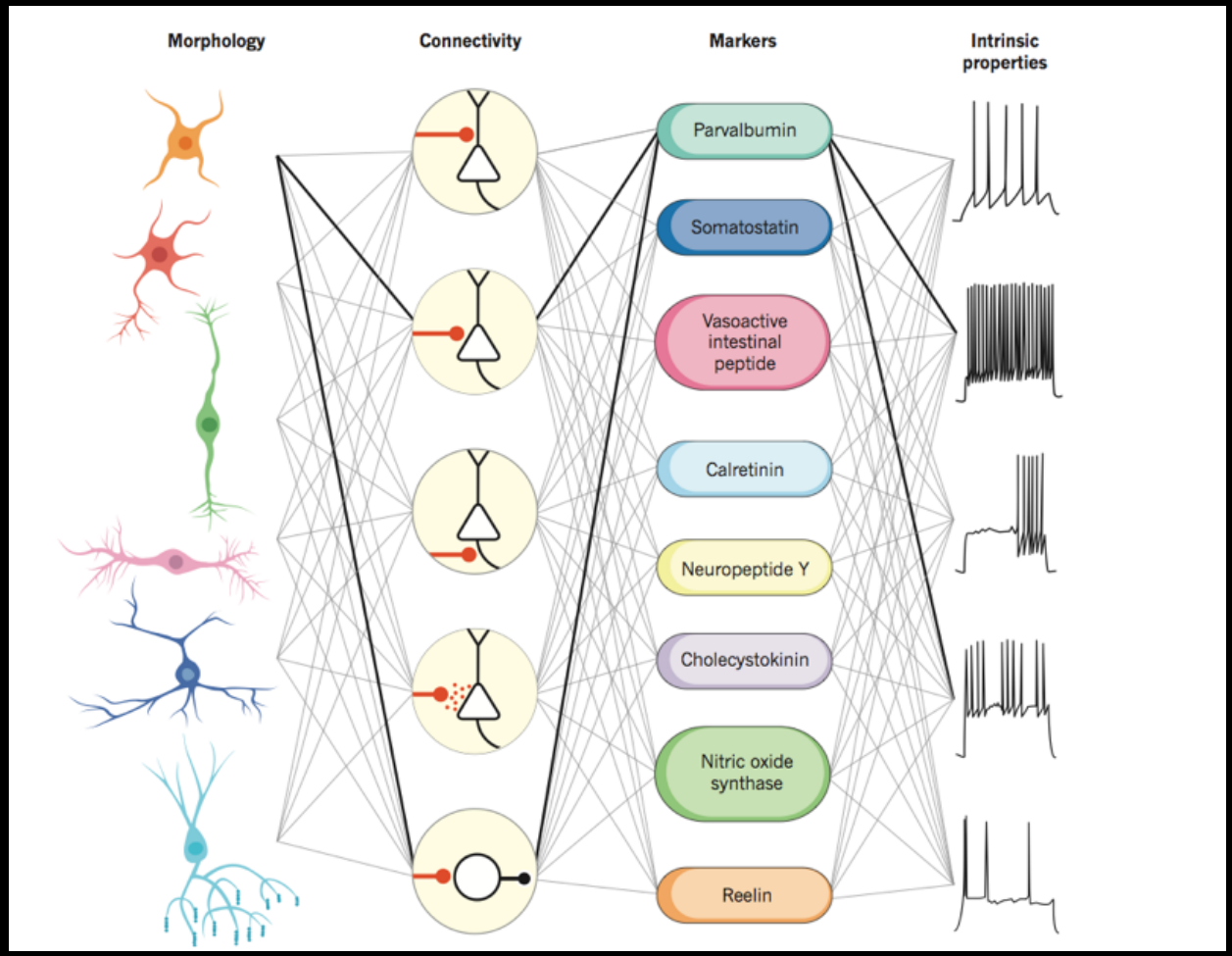




Excitatory cells radiate upward within cortex

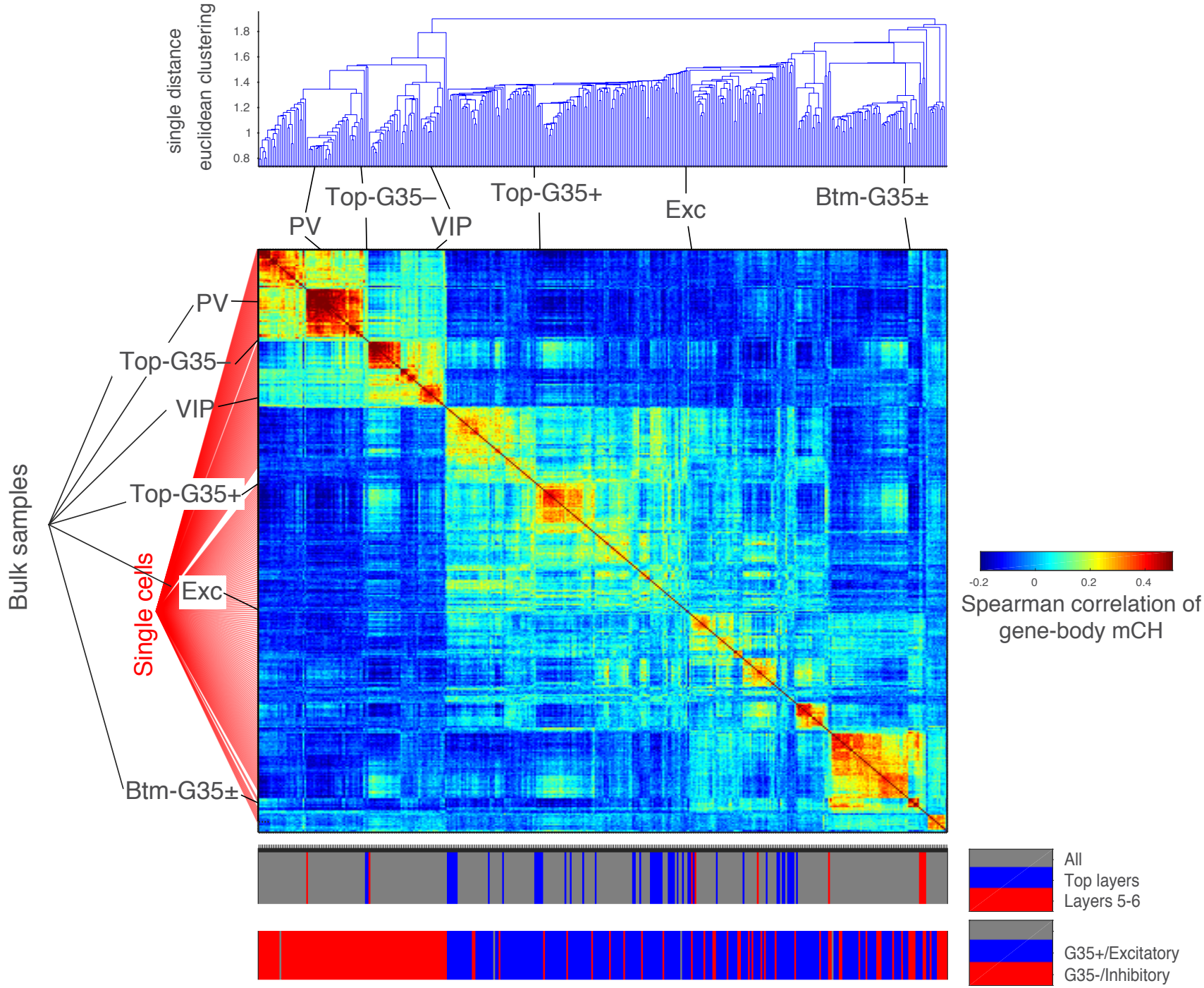
Inhibitory cells migrate to the cortex

What is the DNA methylation landscape in major neuron cell types?



Classification of inhibitory interneuron cell types
Kepecs, A., & Fishell, G. (2014). *Nature*, *505*(7483), 318–326

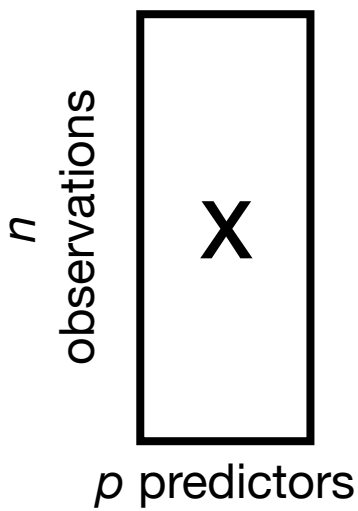
Cluster analysis of 411 single cells + 14 bulk methylomes



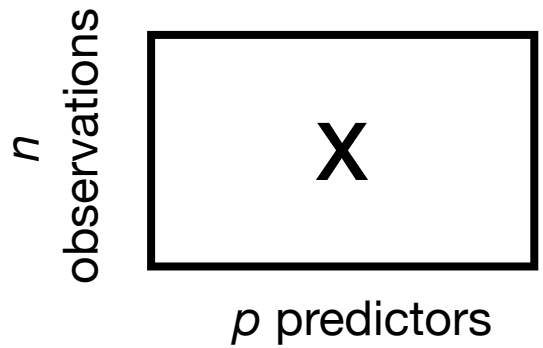
Data dimensionality

- A dataset with *p* “features” (e.g. genes) and *n* “observations (e.g. cells)

- If both *p* and *n* are large (>1,000), it becomes difficult to visualize, analyze and interpret the data



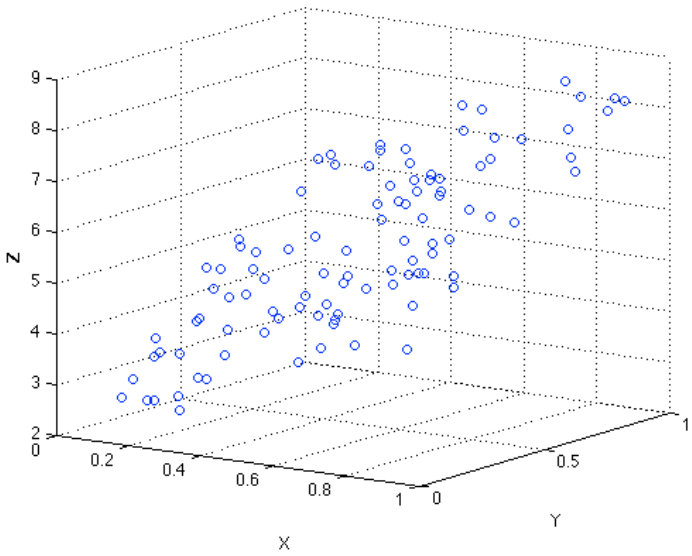
Low-dimensional data, n >> p

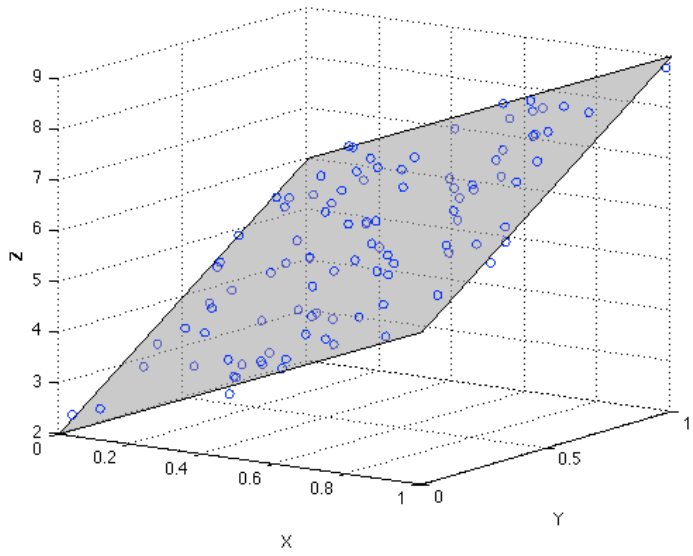


High-dimensional data, n ≤ p

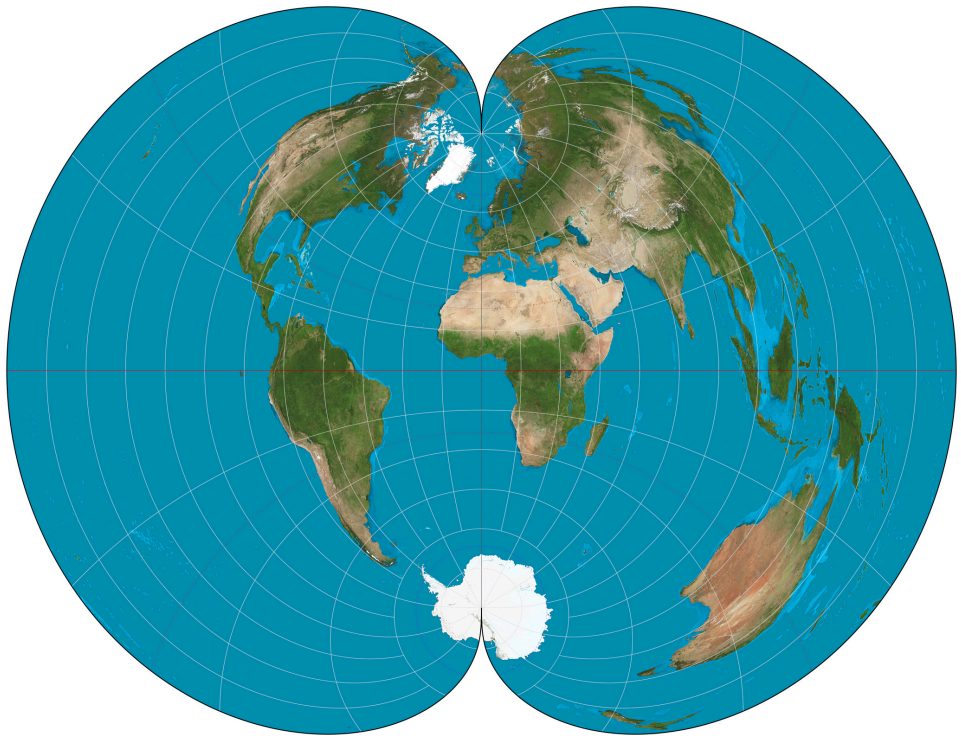
Dimensionality reduction by Principal Components Analysis (PCA)

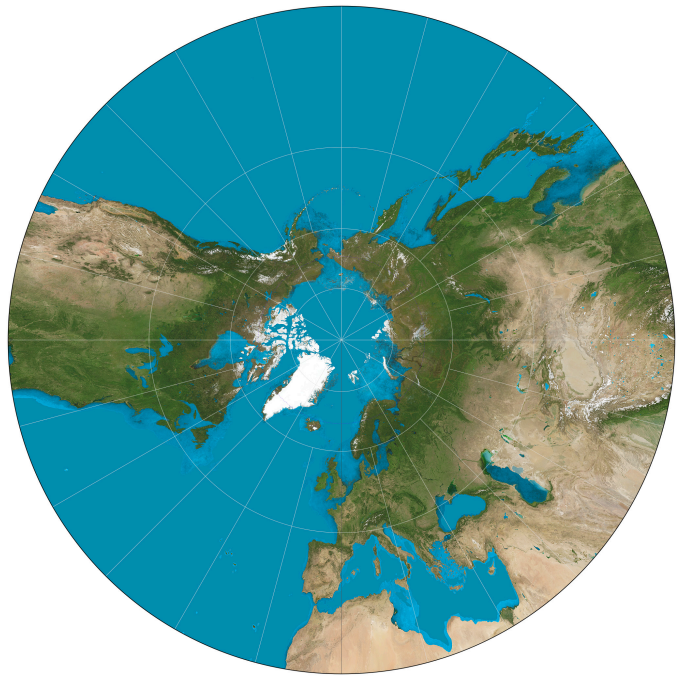
- Principal components analysis (PCA) projects high-dimensional data onto a smaller number of “most interesting” dimensions



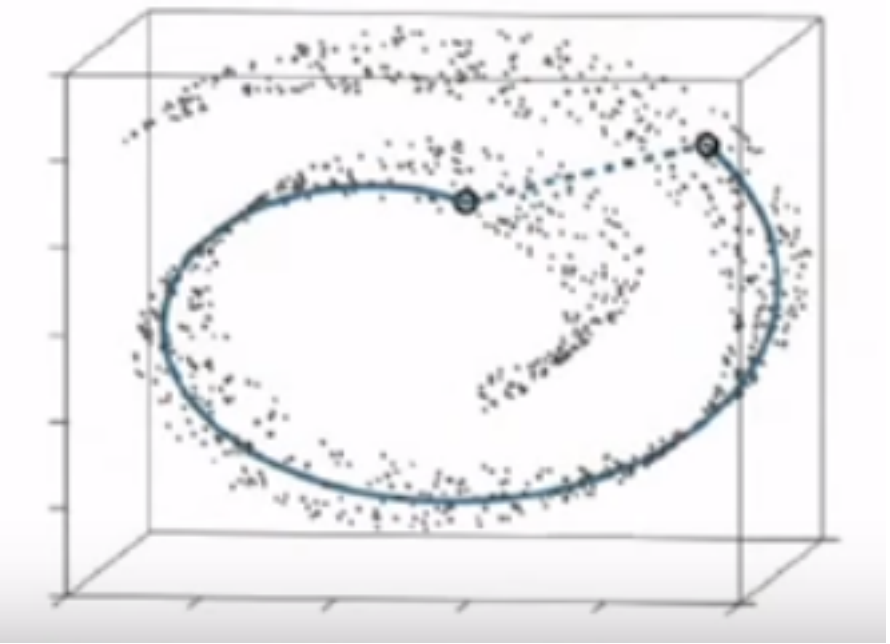


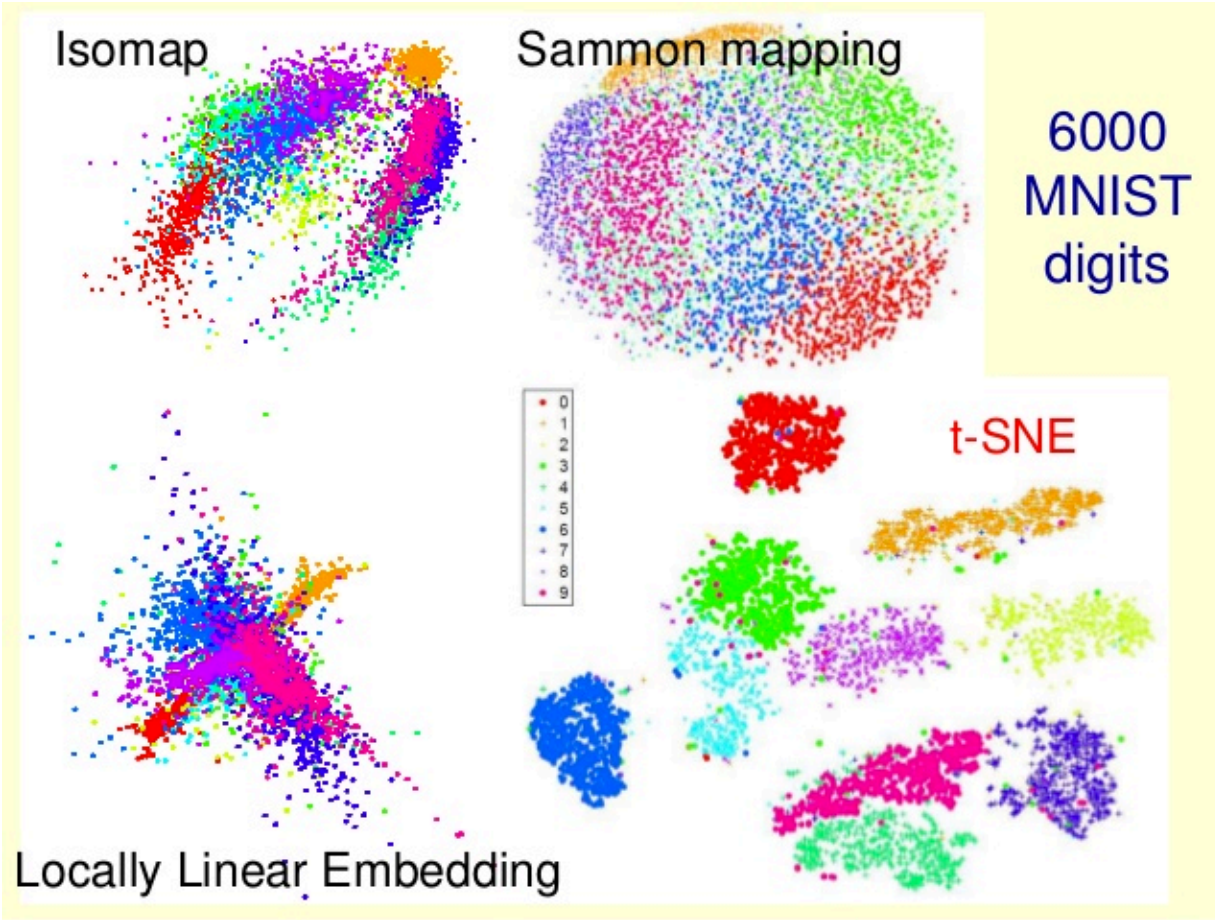
Example: Projection of 3D global geography onto 2D maps





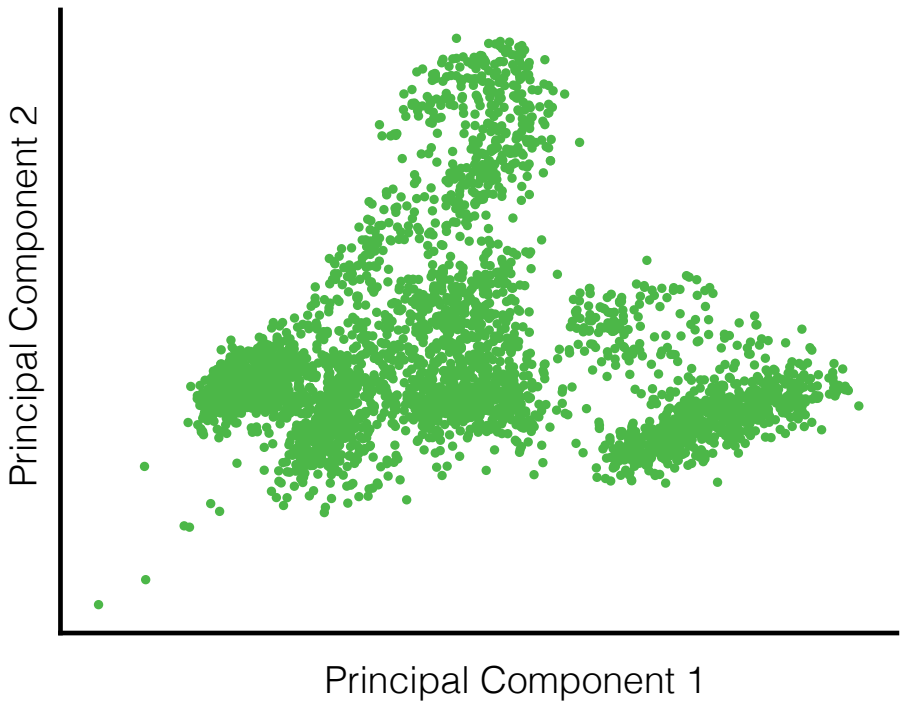
tSNE (t-Stochastic Neighbor Embedding)
Visualizing cells in a high-dimensional space





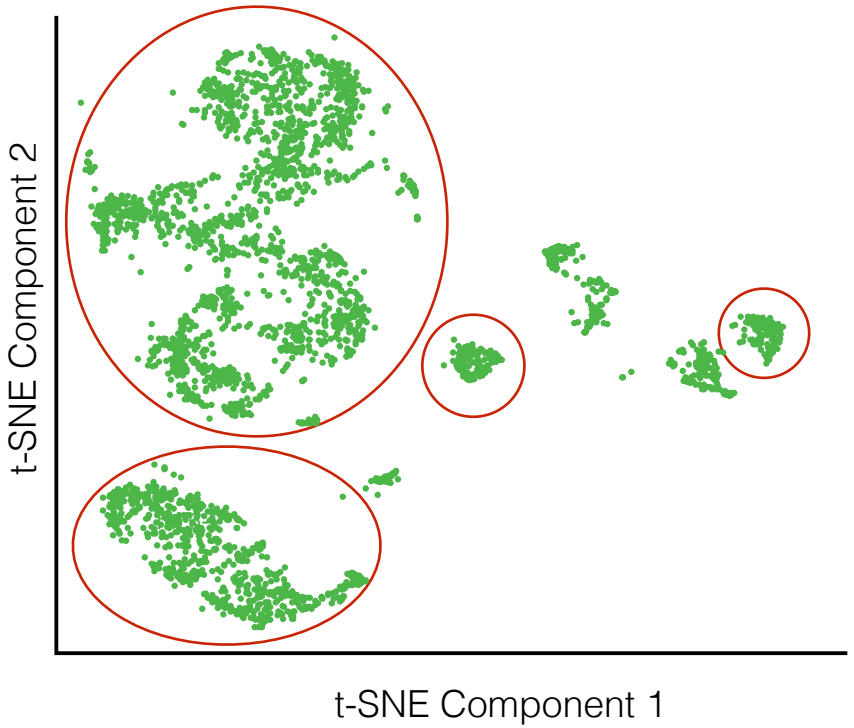
Linear and non-linear dimensional reduction

Principal components analysis

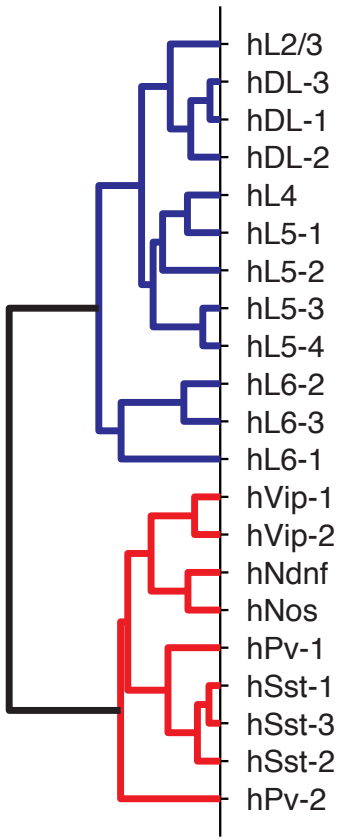


t-Distributed stochastic neighbor embedding

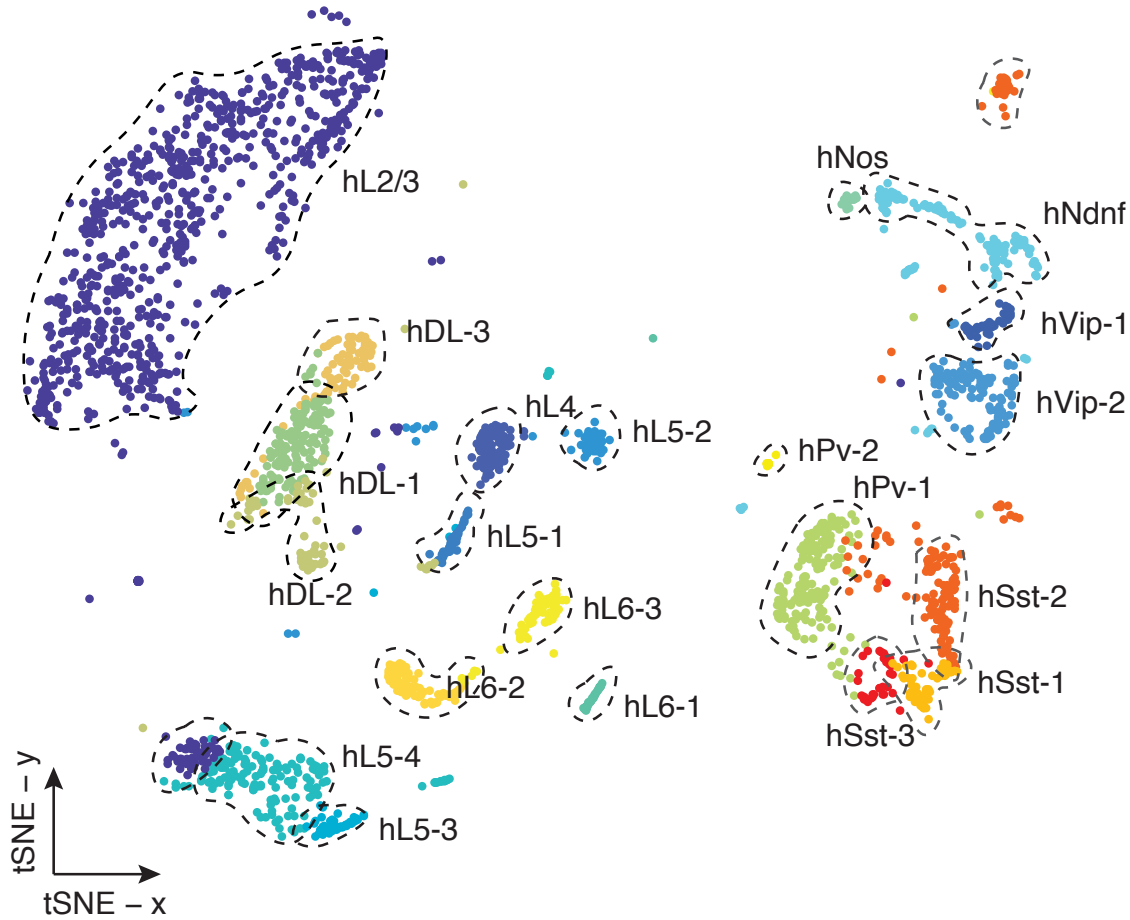
(van der Maaten, Hinton 2008)

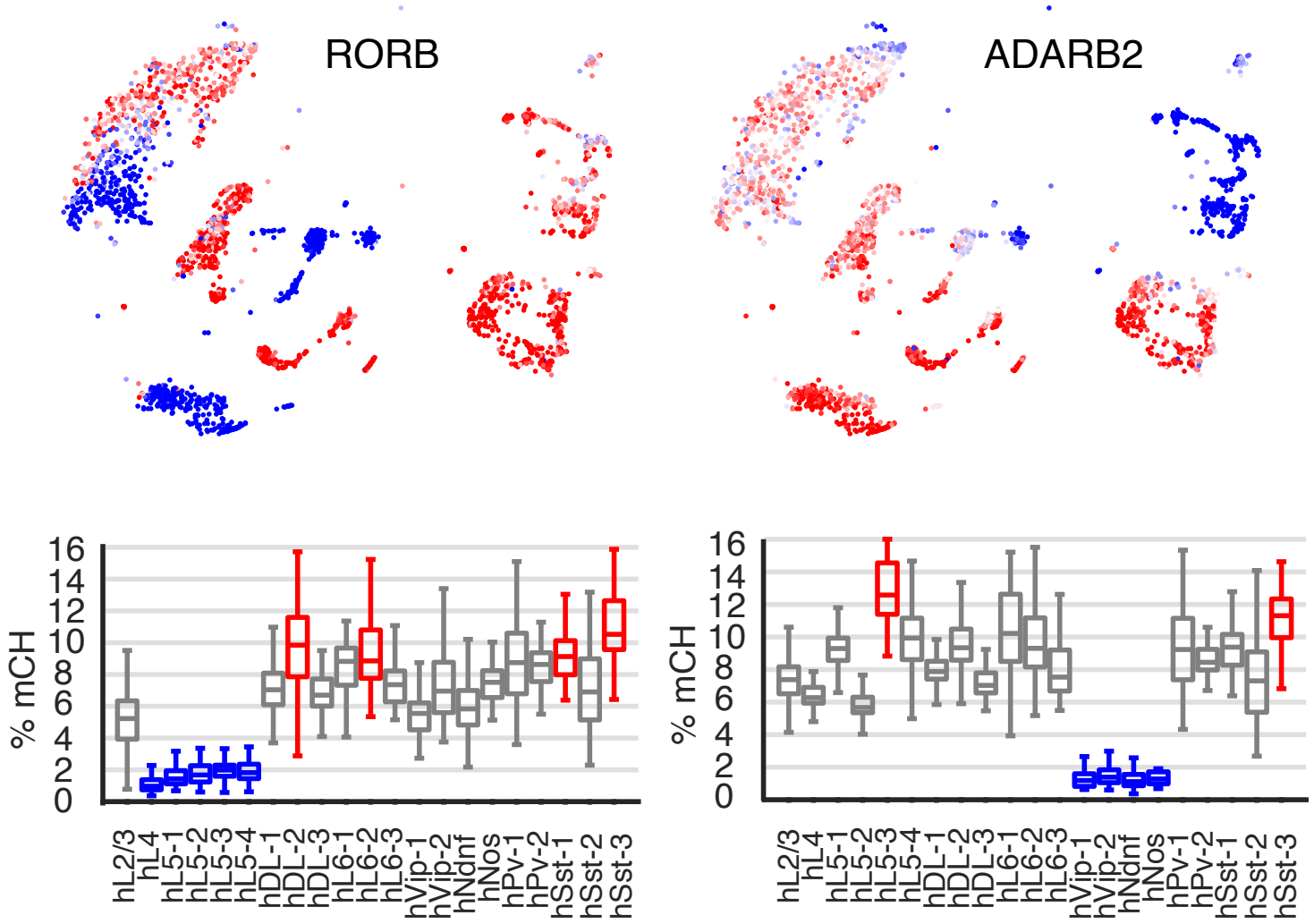


Hierachical clustering
of human neuronal clusters

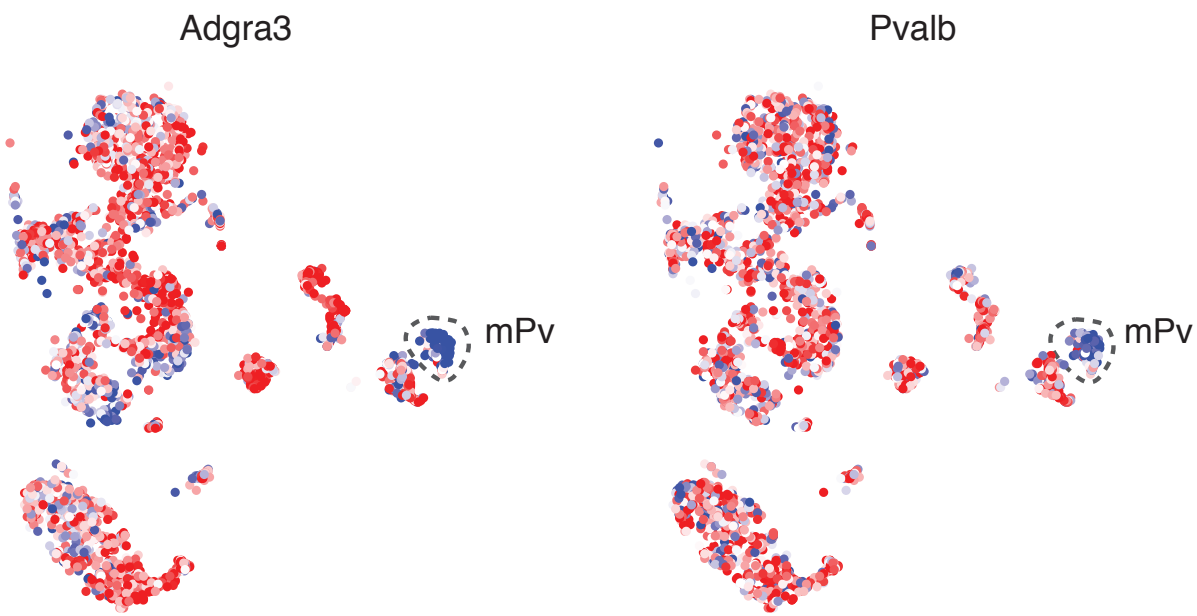


tSNE visualization of single human neurons

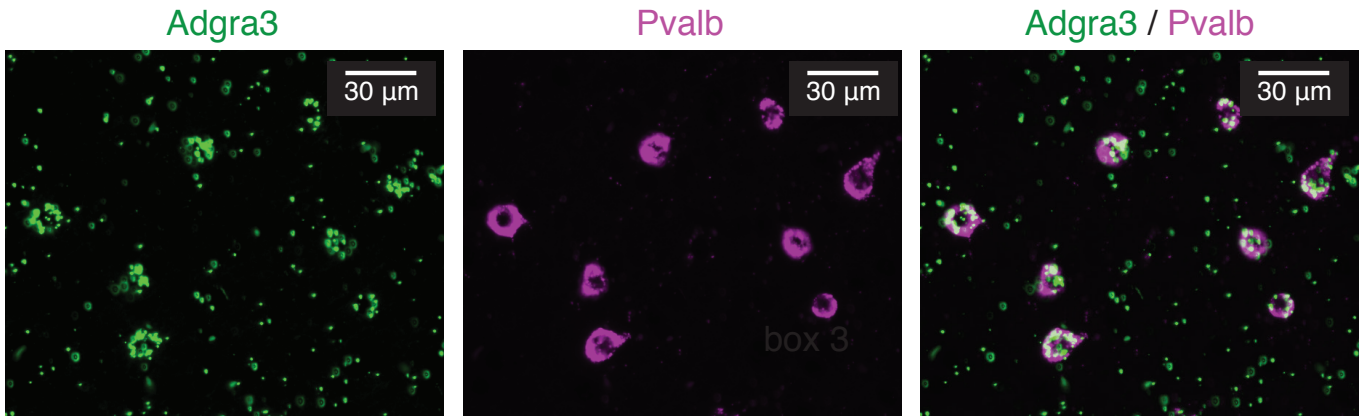


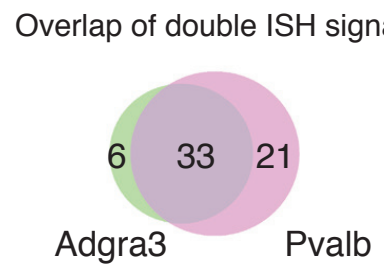

RORB
ADARB2
% mCH
0
2
4
6
8
10
12
14
16
hL2/3
hL4
hL5-1
hL5-2
hL5-3
hL5-4
hDL-1
hDL-2
hDL-3
hL6-1
hL6-2
hL6-3
hVip-1
hVip-2
hNdnf
hNos
hPv-1
hPv-2
hSst-1
hSst-2
hSst-3

Example: Adgra3 is a novel marker of PV+ interneurons



G





How many cell types are there?
Lumpers vs. Splitters

It is good to have hair-splitters & lumpers

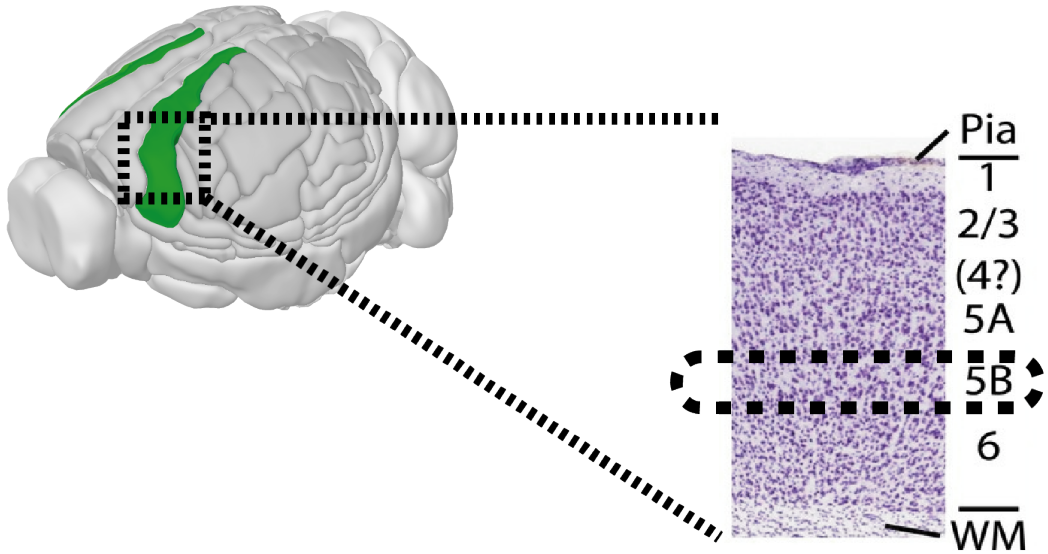
(Darwin, 1857)

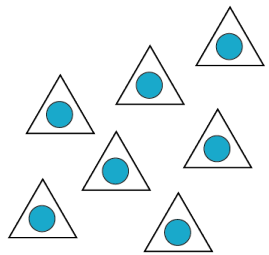
Splitters make very small units – their critics say that if they can tell two animals apart, they place them in different genera ... and if they cannot tell them apart, they place them in different species. ...

Lumpers make large units – their critics say that if a carnivore is neither a dog nor a bear, they call it a cat

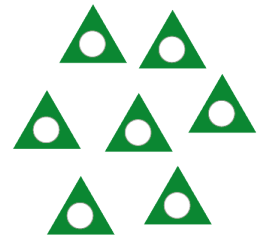
(George Simpson, 1945)

Determining cell types through the integration of multi-modal datasets

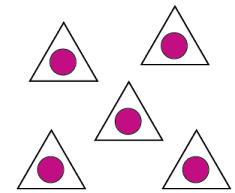




Chromatin accessibility
snATAC-Seq



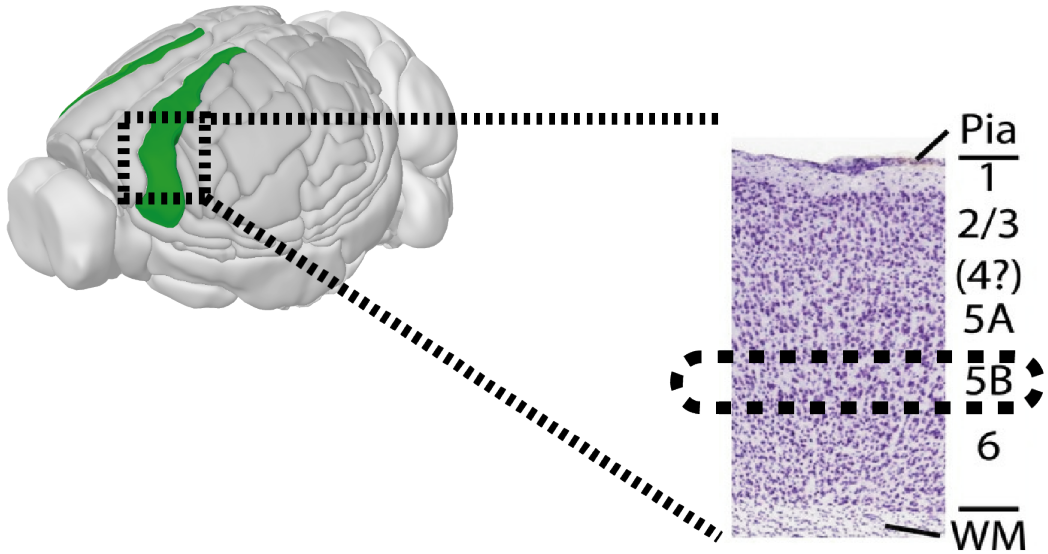
Transcriptome
snRNA/scRNA-Seq

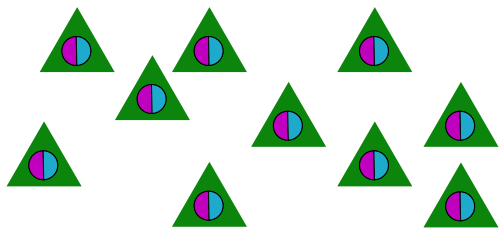


DNA methylation
snmC-Seq

Yamawaki (2014) eLIFE
Brain adapted from Allen Brain Atlas

Determining cell types through the integration of multi-modal datasets



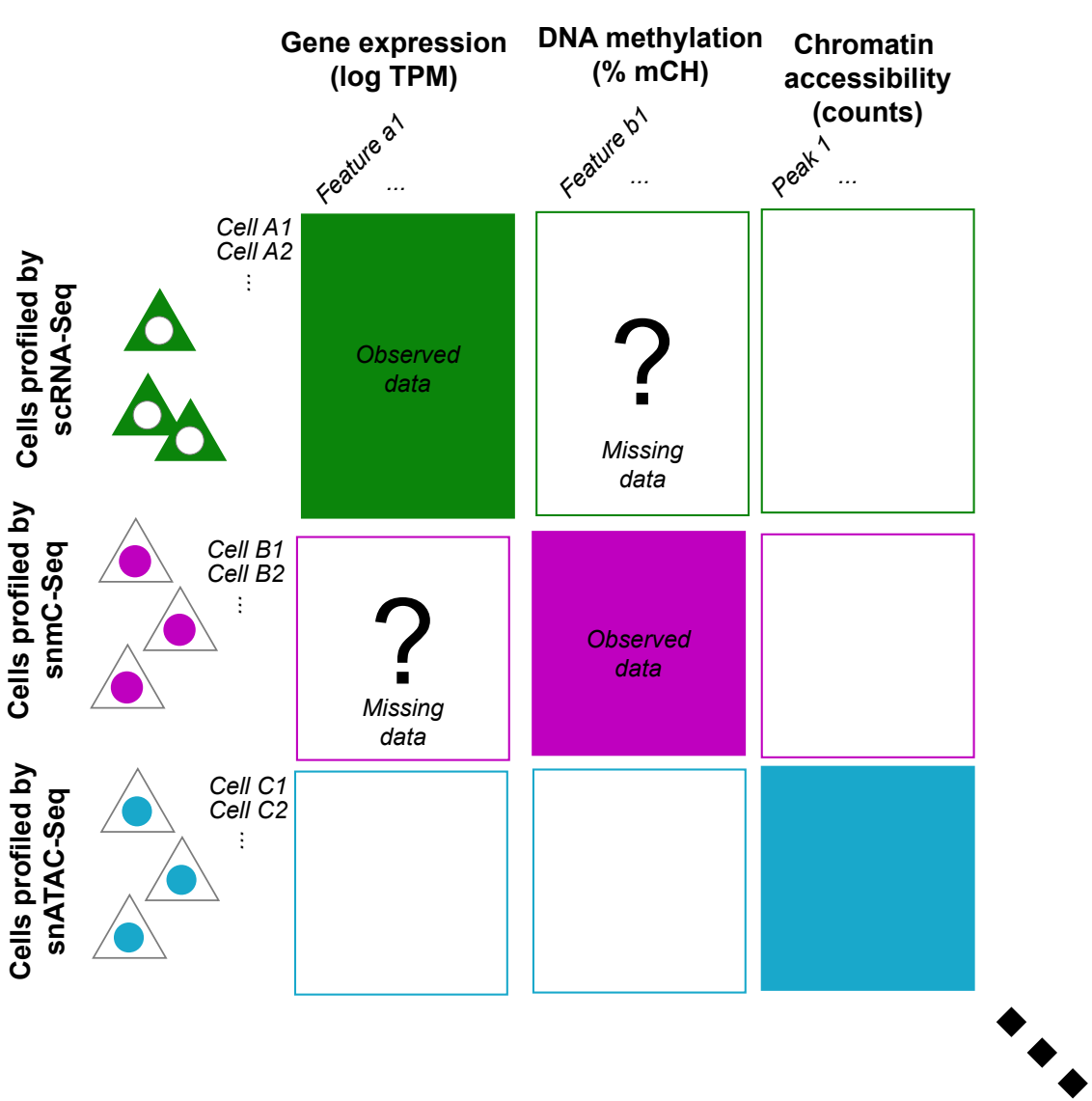


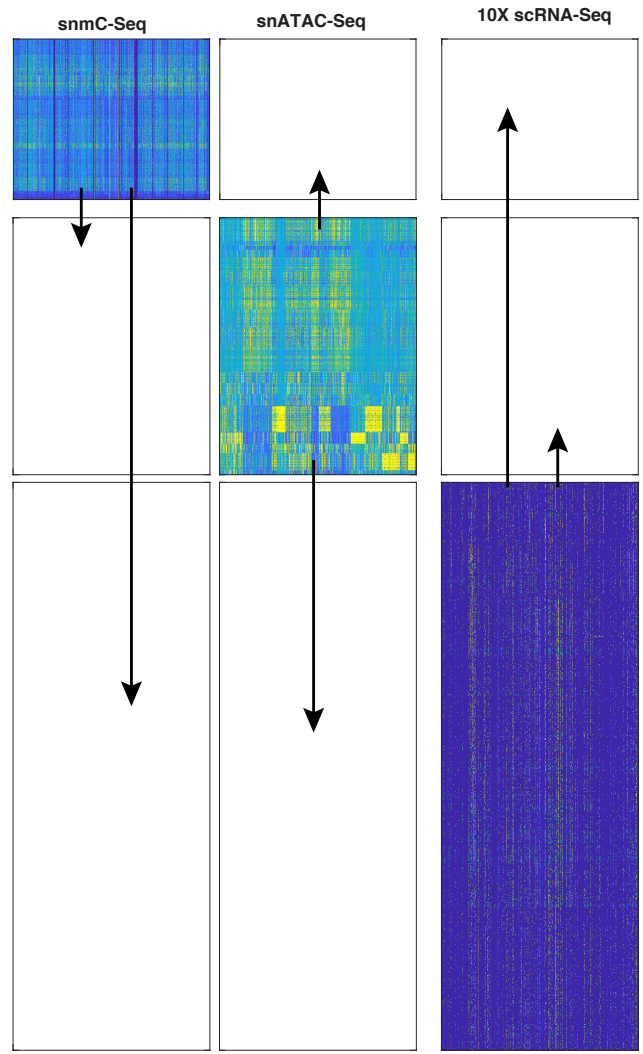
Cell type defined by multiple modalities

- Directly link transcription and epigenetic regulation in the same cell type
- Provide cross-modal validation of cell types predicted from one data modality
- Provides foundation for functional cell types

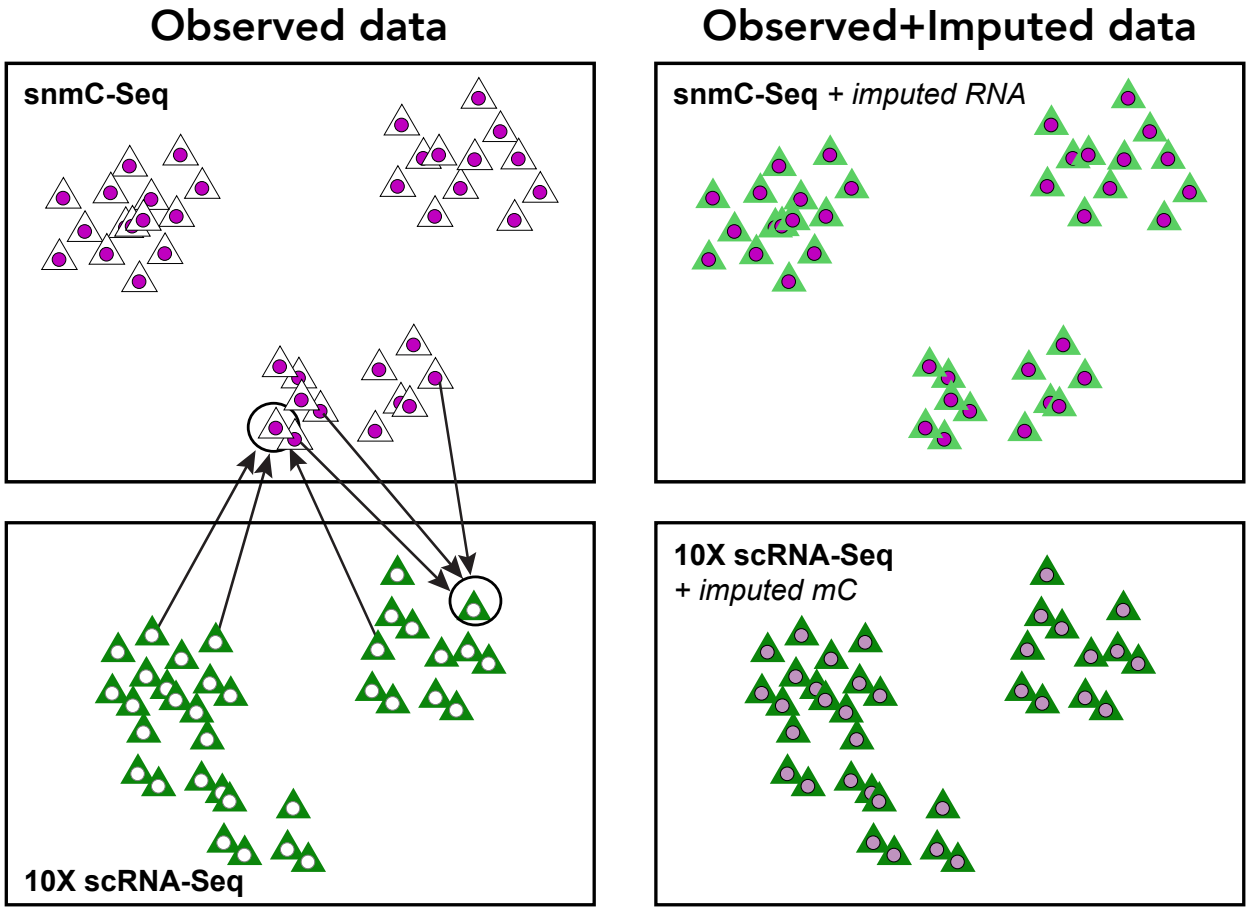
Yamawaki (2014) eLIFE
Brain adapted from Allen Brain Atlas

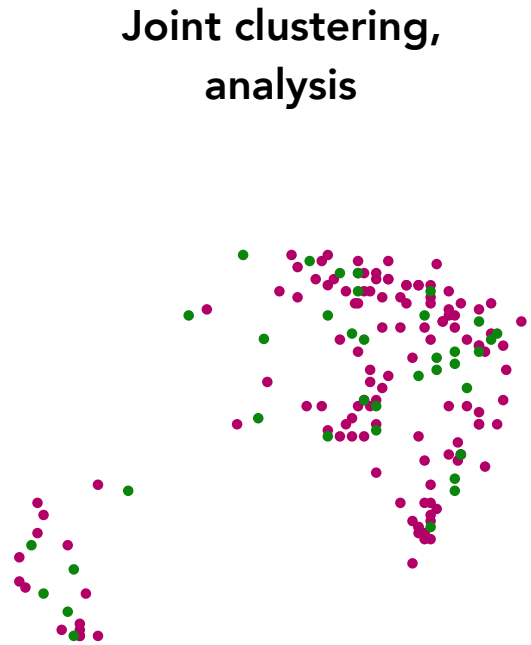
Multi-omics data integration requires imputation



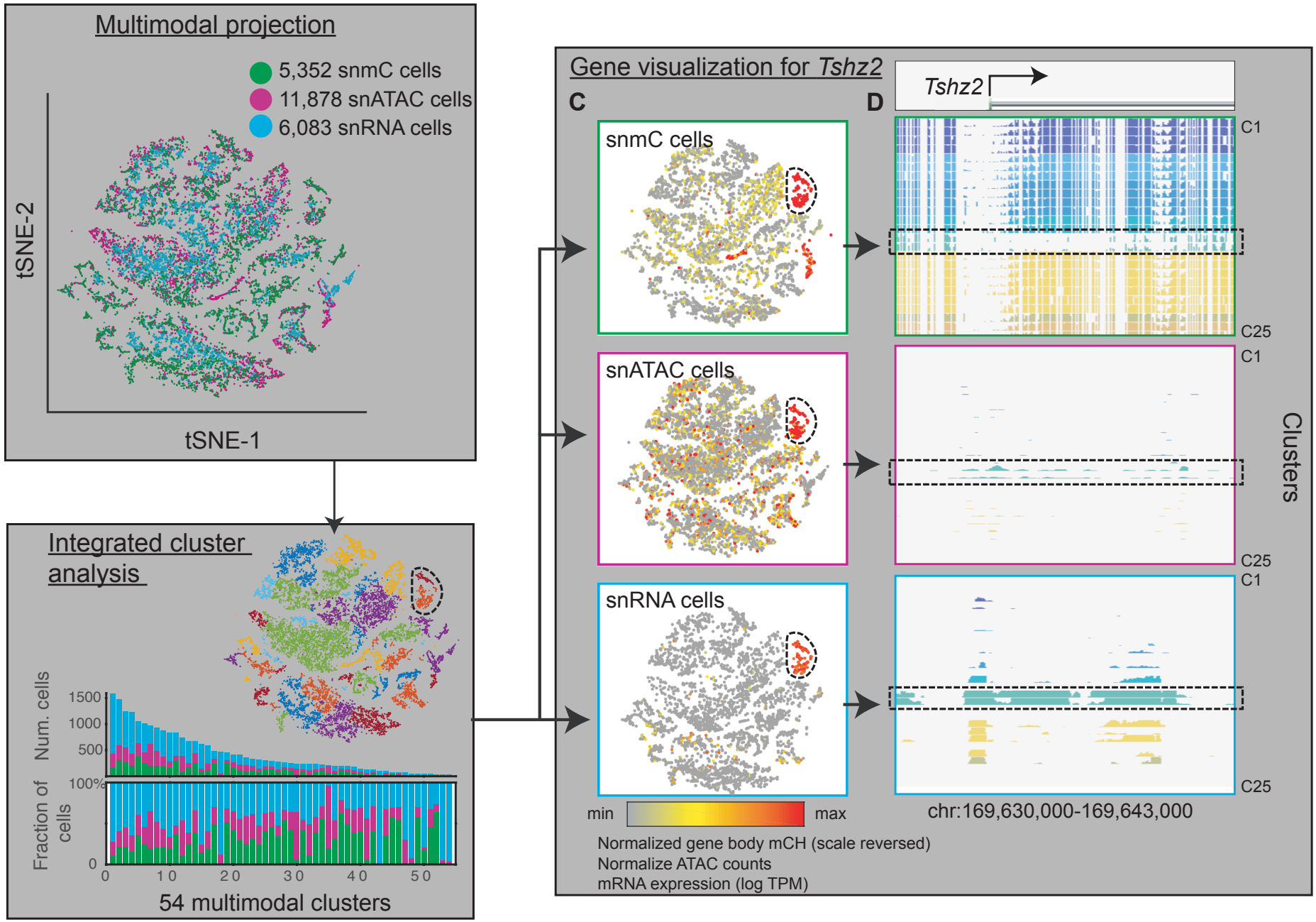


Cross-modality data fusion by *Bigraph Imputation*

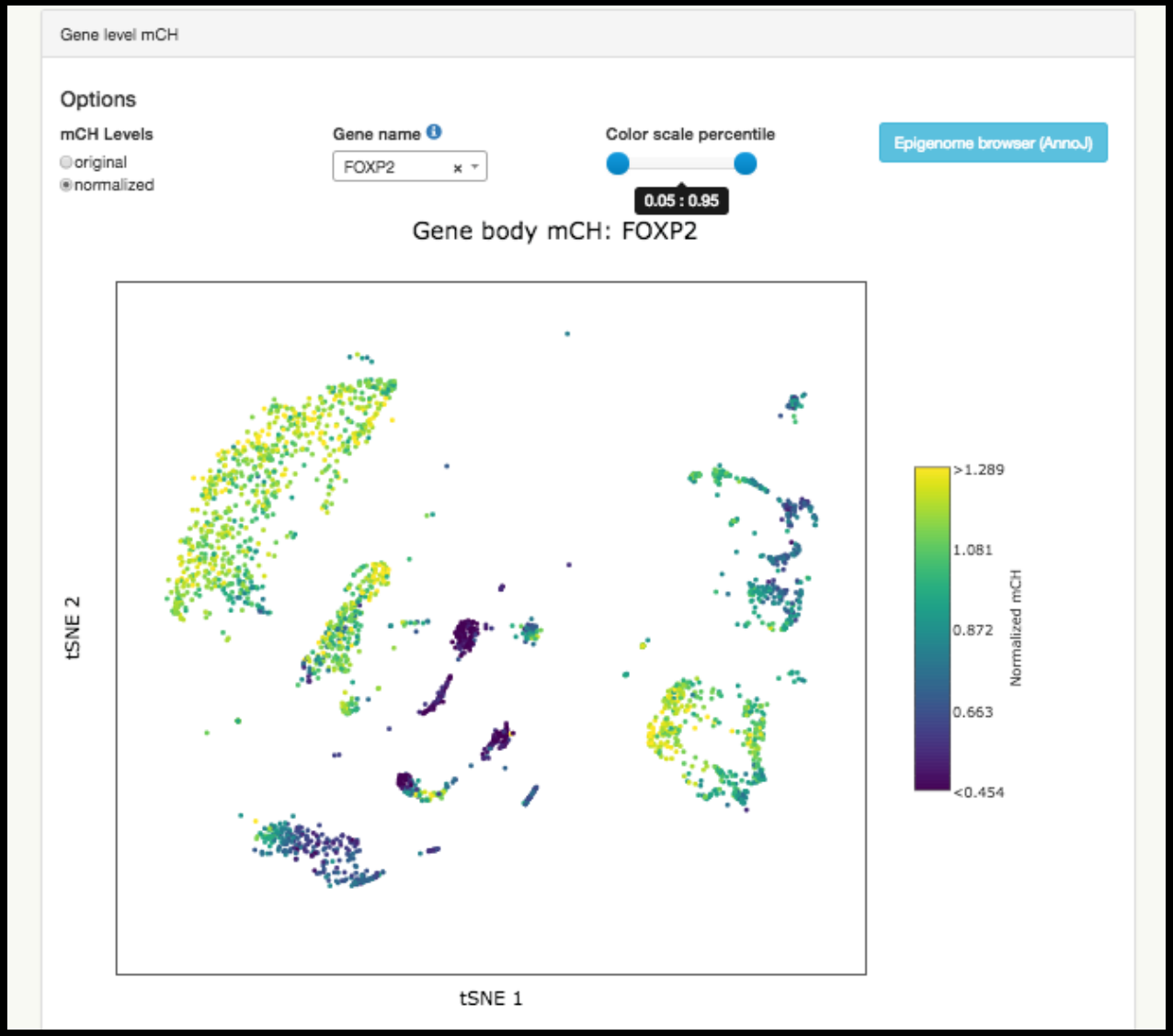




- For each cell in modality A, find K neighbors in modality B
- This requires a *linking assumption*, e.g. low gene body mCH corresponds with high mRNA expression
- Use neighbors to impute missing information for A


Multimodal projection
5,352 snmC cells
11,878 snATAC cells
6,083 snRNA cells
tSNE-2
tSNE-1
Integrated cluster analysis
Num. cells
1500
1000
500
Fraction of cells
100%
0
0
10
20
30
40
50
54 multimodal clusters
Gene visualization for Tshz2
C
D
Tshz2
snmC cells
snATAC cells
snRNA cells
C1
C25
C1
C25
C1
C25
Clusters
min
max
chr:169,630,000-169,643,000
Normalized gene body mCH (scale reversed)
Normalize ATAC counts
mRNA expression (log TPM)

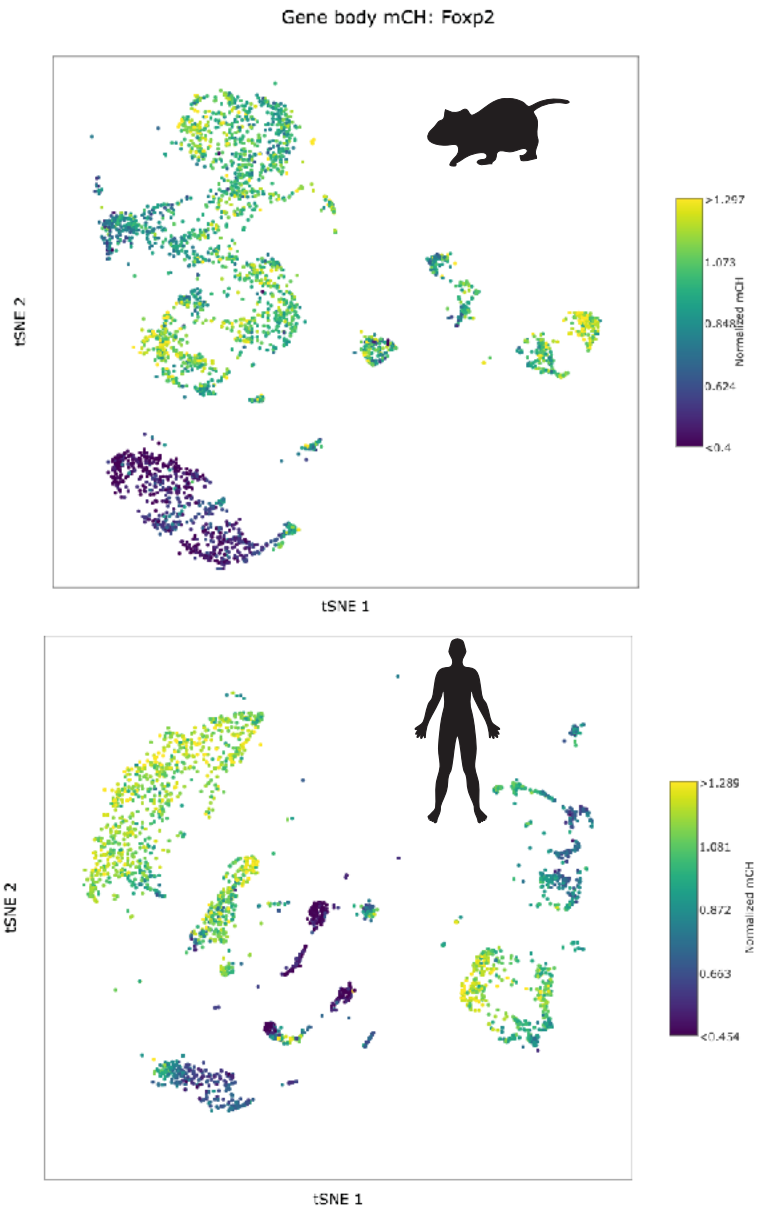
Exploring cell type specific DNA methylation patterns with interactive data visualizations (http://brainome.org)

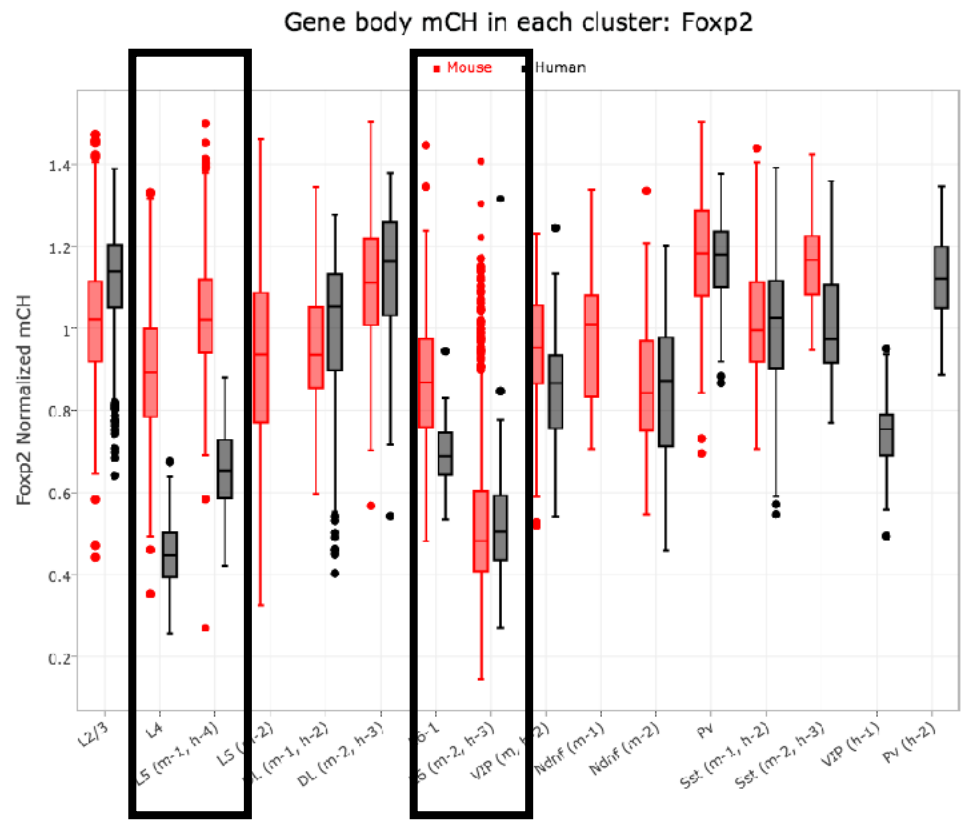


FOXP2:
A “language gene”?

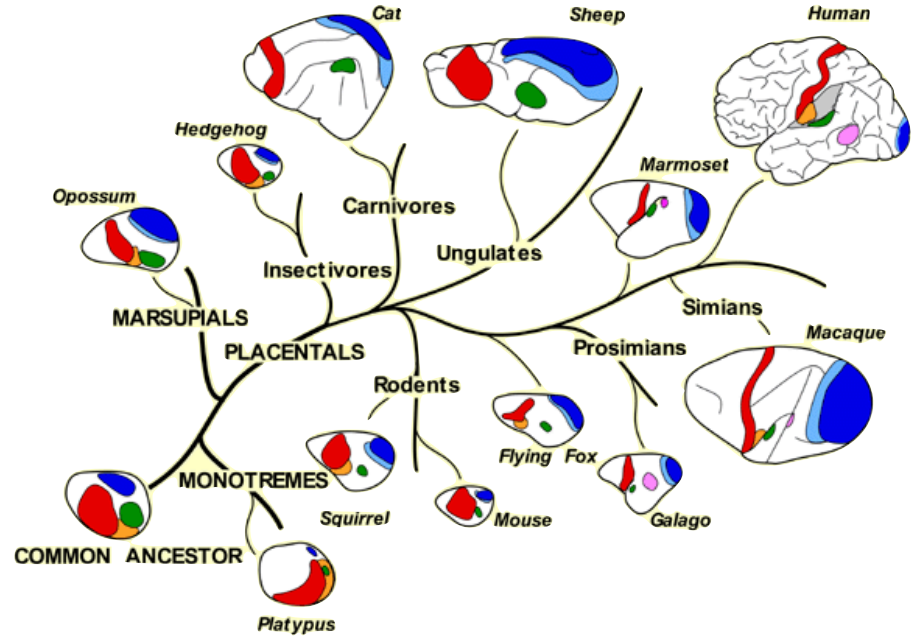
- Mutations in FOXP2 linked to language disorders (verbal dyspraxia)
- Also linked with vocalization in songbirds

FOXP2 (associated with language) is expressed in different cortical layers in mouse and human frontal cortex

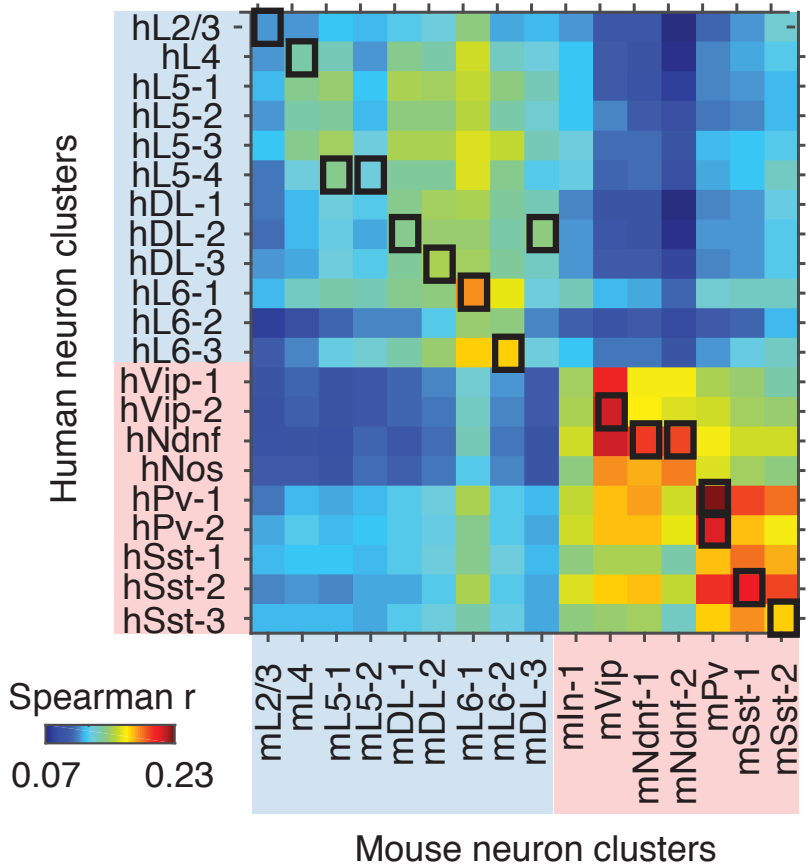




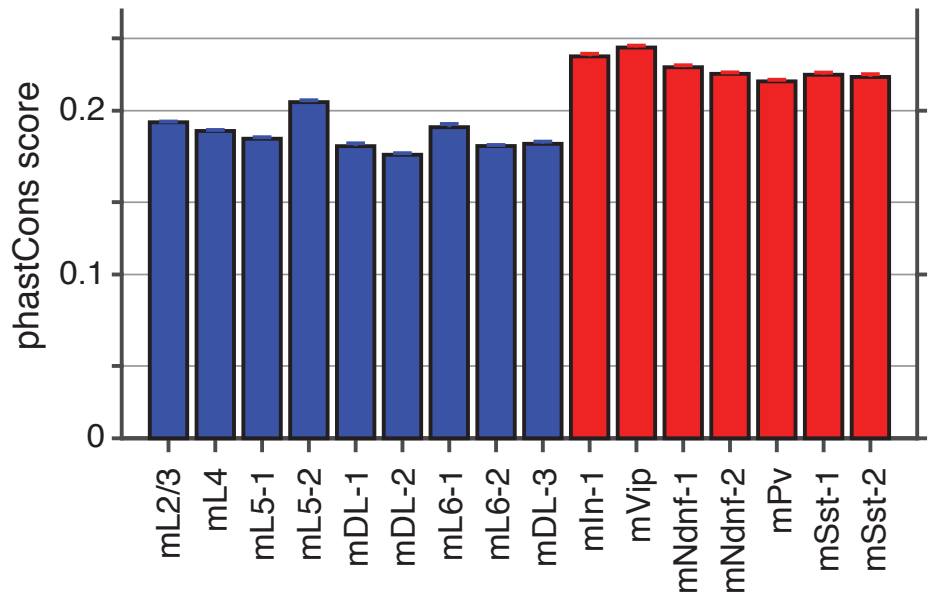
~60M years of evolution led to greater divergence in excitatory compared with inhibitory neuron epigenomes



Cross-species correlation of mCG at CG-DMRs

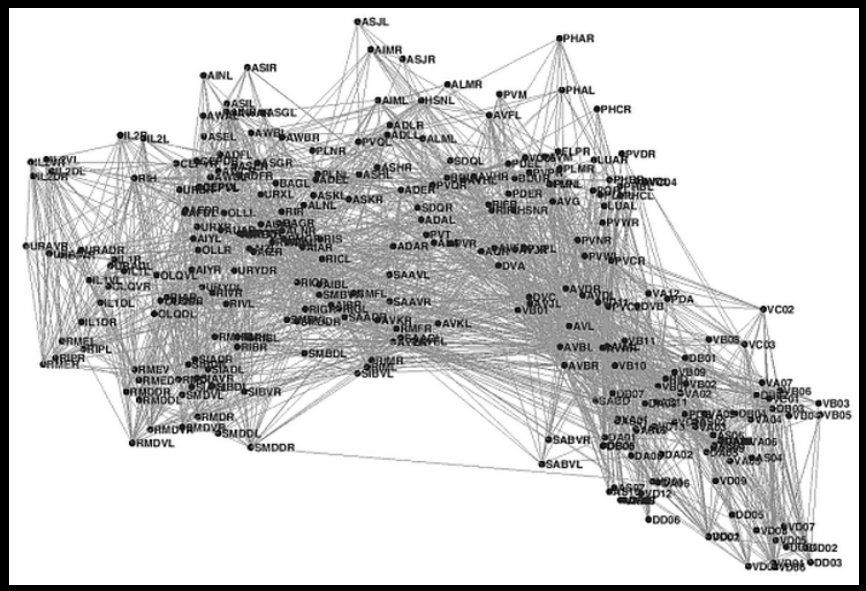


Sequence conservation of excitatory and inhibitory neuron CG-DMRs



Complex biological networks: Genes and Brains

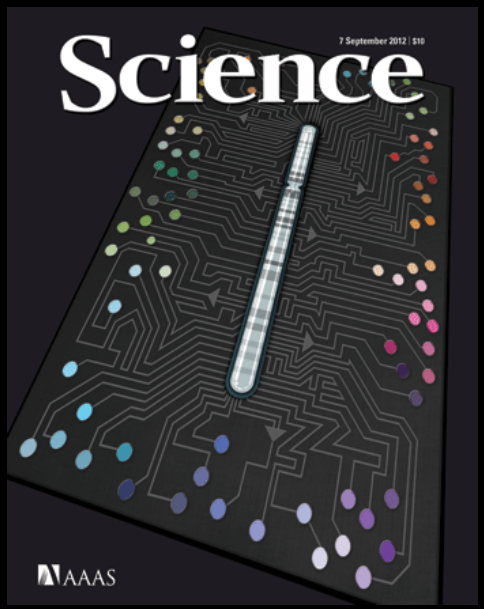
Brains (neural networks)



C. Elegans connectome (1986)

?

Genomes (gene networks)



Encyclopedia of DNA Elements (ENCODE) (2012)

✓	Encode and store innate information	✓
✓	Encode and store learned information	✓
✓	Transmit information	✓
✓	Enable complex, recurrent interactions	✓

Take-home message

The methylation status at ~1 billion cytosines in the genome is potentially an information-rich, stable yet flexible substrate for information storage/processing

Open questions:

- *Functional role of mCH?*
- *Differences between 100s of neuronal subtypes?*
- *Impact of experience/learning?*
- *Disruption in neuropsychiatric disorders?*

Acknowledgments

UCSD Cognitive Science

- Chris Keown
- Andrew Schork

Salk Computational Neurobiology Lab

- Terry Sejnowski
- Margarita Behrens
- Nick Johnson
- Clare Puddifoot

Salk Genomic Analysis Lab

- Joe Ecker
- Ryan Lister (now U. Western Australia)
- Chongyuan He
- Yupeng He
- Matt Schultz

Johns Hopkins

- Jeremy Nathans (Johns Hopkins)
- Alisa Mo (Johns Hopkins)

HHMI - Janelia Farms

- Sean Eddy (Janelia)
- Fred Davis (Janelia)

Funding

NIH NINDS K99/R00